

Figure 1

CGAGAGACGACAGAAGGTTACGGCTGCGAGAAGACGACAGAAGGGTCCAGAAAAA
GGAAAGTGCTGGAGGGGAGTGGGGACAAAAGCAGCGACCAAGTGAATGTCACTTC
AGTGACTGAGGCCAGGCAAAACGCGCGGGAAGGATTTTGTGTAGCTTGGGACCCTT
TCATAGACACTGATGACACGTTTACGCAAAATAGAAATTTGAGGAGAAACGCCTGG
GCCTTCGGAAAGGAGTGATTGATTAGTACTTGCAAGTTTAGGTGACTTTAAGGAGAA
CTAACTAATGTATACTATTGAGGGAGGAGGAAGAGCATTACAGAGTTTCCAGCAGC
AGCAGGAAAGCTTTGGTTAATTTGGAAATGGATGATAGCATTAAAATAACAGAAGC
GCCTCCAGGTCTCTGAAGCTTCAGTCCCCCAGCTGAAAGCCAGAAAAGACTAAGCC
CACTAAGCCTTTTGATCCCTTTGGAAGCAAAGAACTTTCCTTCCCTGGGGTGAAGAC
TCTCCTCAGAAGATTTCTGTCTCTGCCTATGTTACAAGAGGAATCAAAACCAAGAC
AGAAGAGCTCAGGATGCAGGTGAGAGGCAGGGAAGTCAGCGGCTTGTTGATCTCCC
TCACTGCTGTCTGCCTGGTGGTCACCCCTGGGAGCAGGGCCTGTCCTCGCCGCTGTG
CCTGCTATGTGCCACAGAGGTGCACTGTACATTTTCGGTACCTGACCTCCATCCCAG
ATGGCATCCCGGCCAATGTGGAACGAATAAATTTAGGATATAACAGCCTTACTAGAT
TGACAGAAAACGACTTTGATGGCCTGAGCAAACCTGGAGTTACTCATGCTGCACAGT
AATGGCATTACAGAGTCAGTGACAAGACCTTCTCGGGCTTGACAGTCCTTGACGGTC
TAAAAATGAGCTATAACAAAGTCCAAATCATTTCGGAAGGATACTTTCTACGGACTC
GGGAGCTTGGTCCGGTTGCACCTGGATCACAACAACATTGAATTCATCAACCCTGAG
GCCTTTTATGGACTTACCTCGCTCCGCTTGGTACATTTAGAAGGAAACCGGCTCACA
AAGCTCCATCCAGACACATTTGTCTCATTAAAGCTATCTCCAGATATTTAAAACCTCTT
TCATTAAGTACCTGTTCTTGTCTGATAACTTCCTGACCTCCCTCCCAAAGAAATGGT
CTCCTACATGCCAAACCTAGAAAGCCTGTATTTGCATGGAAACCCATGGACCTGTGA
CTGCCATTTAAAGTGGTTGTCTGAGTGGATGCAGGGAAACCCAGATATAATAAAAT
GCAAGAAAGACAGAAGCTCTTCCAGTCCTCAGCAATGTCCCCTTTGCATGAACCCCA
GGATCTCTAAAGGCAGACCCTTTGCTATGGTACCATCTGGAGCTTTCCTATGTACAA
AGCCAACCATTGATCCATCACTGAAGTCAAAGAGCCTGGTTACTCAGGAGGACAAT
GGATCTGCCTCCACCTCACCTCAAGATTTTCATAGAACCCTTTGGCTCCTTGTCTTTGA
ACATGACANANNTNTCTGGAAATAAGGCCGACATGGTCTGTAGTATCCAAAAGCCA
TCAAGGACATCACCAACTGCATTCACTGAAGAAAATGACTACATCATGCTAAATGC
GTCATTTTCCACAAATCTTGTGTGCAGTGTAGATTATAATCACATCCAGCCAGTGTG
GCAACTTCTGGCTTTATACAGTGACTCTCCTCTGATACTAGAAAGGAAGCCCCAGCT
TACCGAGACTCCTTCACTGTCTTCTAGATATAAACAGGTGGCTCTTAGGCCTGAAGA
CATTTTTACCAGCATAGAGGCTGATGTCAGAGCAGACCCTTTTTGGTTCCAACAAGA
AAAAATTGTCTTGCAGCTGAACAGAACTGCCACCACACTTAGCACATTACAGATCCA
GTTTTCCACTGATGCTCAAATCGCTTTACCAAGGGCGGAGATGAGAGCGGAGAGAC
TCAAATGGACCATGATCCTGATGATGAACAATCCCAAACCTGGAACGCACTGTCCTGG
TTGGCGGCACTATTGCCCTGAGCTGTCCAGGCAAAGGCGACCCTTCACCTCACTTGG
AATGGCTTCTAGCTGATGGGAGTAAAGTGAGAGCCCCTTACGTTAGCGAGGATGGG

CGAATCCTAATAGACAAAAATGGGAAGTTGGAAGTGCAGATGGCTGACAGCTTTGA
TGCAGGTCTTTACCACTGCATAAGCACCAATGATGCAGATGCGGATGTTCTCACATA
CAGGATAACTGTGGTAGAGCCCTATGGAGAAAGCACACATGACAGTGGAGTCCAGC
ACACAGTGGTTACGGGTGAGACGCTCGACCTTCCATGCCTTTCCACGGGTGTTCCAG
ATGCTTCTATTAGCTGGATTCTTCCAGGGAACACTGTGTTCTCTCAGCCATCAAGAG
ACAGGCAAATTCTTAACAATGGGACCTTAAGAATATTACAGGTTACGCCAAAAGAT
CAAGGTCATTACCAATGTGTGGCTGCCAACCATCAGGGGCCGACTTTTCCAGTTTT
AAAGTTTCAGTTCAAAAGAAAGGCCAAAGGATGGTTGAGCATGACAGGGAGGCAG
GTGGATCTGGACTTGGAGAACCCAACTCCAGTGTTTCCCTTAAGCAGCCAGCATCTT
TGAAACTCTCTGCATCAGCTTTGACAGGGTCAGAGGCTGGAAAACAAGTCTCCGGTG
TACATAGGAAGAACAACATAGAGACTTAATACATCGGCGGCGTGGGGAATTCCACG
CTCCGGCGATTACAGGGAGCATAGGAGGCAGCTCCCTCTCTCTGCTCGGAGAATTGAC
CCGCAACGCTGGGCAGCACTTCTAGAAAAAGCCAAAAAGAATTCTGTGCCAAAAAA
GCAAGAAAATACCACAGTAAAGCCAGTGCCACTGGCTGTTCCCCTCGTGGAAGTCA
CTGACGAGGAAAAGGATGCCTCTGGCATGATTCTCTCCAGATGAAGAATTCATGGTTC
TGAAAATAAGGCTTCTGGTGTCCAGGAAGGTCACCAACTGCTGACTCTGGACCAG
TAAATCATGGTTTTATGACGAGTATAGCTTCTGGCACAGAAGTCTCAACTGTGAATC
CACAAACACTACAATCTGAGCACCTTCCTGATTTCAAATTATTTAGTGTAACAAACG
GTACAGCTGTGACAAAGAGTATGAACCCATCCATAGCAAGCAAAATAGAAGATACA
ACCAACCAAAACCCAATCATTATCTTTCCATCAGTAGCTGAAATTCGAGATTCTGCT
CAGGCAGGAAGAGCATCTTCCCAAAGTGCACACCCTGTAACAGGGGGGAAACATGGC
TACCTATGGCCATACCAACACATATAGTAGCTTTACCAGCAAAGCCAGTACAGTCTT
GCAGCCAATAAATCCAACAGAAAGTTATGGACCTCAGATACCTATTACAGGAGTCA
GCAGACCTAGCAGTAGTGACATCTTCTCACACTACTGCAGACCCTAGCTTCTCCA
GTCACCCTTCAGGTTACACACCACTGCCTCGTCTTTATTTACATTCTAGAAACAA
CAATACAGGTAAGTTCCCCTTGTCAGGCACTTGGGAAGAGAGAGGACAATTTGGA
GCAGAGGGGAGAGTTAAAAACCCACATAGAACCCCAAGTTCTCCGACGGCATAGACAC
AGGACTGTGAGGCCAGCAATCAAGGGACCTGCTAACAAAAATGTGAGCCAAGTTCC
AGCCACAGAGTACCCTGGGATGTGCCACACATGTCCTTCCGCAGAGGGGGCTCACAG
TGGCTACTGCAGCACTGTCAGTTCCAAGTTCATCCACAGTGCCCTCCCCAAAACTA
ATAATGTTGGGGTCATAGCAGAAGAGTCTACCACTGTGGTCAAGAAACCACTGTTAC
TATTTAAGGACAAACAAAATGTAGATATTGAGATAATAACAACCACTACAAAATAT
TCCGGAGGGGAAAGTAACCACGTGATTCCTACGGAAGCAAGCATGACTTCTGCTCC
AACATCTGTATCCCTGGGGAAATCTCCTGTAGACAATAGTGGTCACCTGAGCATGCC
TGGGACCATCCAACTGGGAAAGATTGAGTGGAAACAACACCACTTCCCAGCCCCC
TCAGCACACCCTCAATACCAACAAGCACAAAATTCTCAAAGAGGAAAACTCCCTTG
CACCAGATCTTTGTAAATAACCAGAAGAAGGAGGGGATGTTAAAGAATCCATATCA
ATTCGGTTTACAAAAGAACCCAGCCGCAAAGCTTCCCAAATAGCTCCTCTTTTACC
CACAGGTCAGAGTTCCCCCTCAGATTCTACAACCTCTTTGACAAGTCCGCCACCAGC
TCTGTCTACAACAATGGCTGCCACTCAGAACAAGGGCACTGAAGTAGTATCAGGTG

CCAGAAGTCTCTCAGCAGGGAAGAAGCAGCCCTTCACCAACTCCTCTCCAGTGCTTC
CTAGCACCATAAGCAAGAGATCTAATACATTAAACTTCTTGTCAACGGAAACCCCCA
CAGTGACAAGTCCTACTGCTACTGCATCTGTCAATTATGTCTGAAACCCAACGAACAA
GATCCAAAGAAGCAAAAGACCAAATAAAGGGGCCTCGGAAGAACAGAAACAACGC
AAACACCACCCCCAGGCAGGTTTCTGGCTATAGTGCATACTCAGCTCTAACAACAGC
TGATACCCCCTTGGCTTTTCAGTCATTCCCCACGACAAGATGATGGTGGAATGTAAG
TGCAGTTGCTTATCACTCAACAACCTCTCTTCTGGCCATAACTGAACTGTTTGAGAA
GTACACCCAGACTTTGGGAAATACAACAGCTTTGGAAACAACGTTGTTGAGCAAAT
CACAGGAGAGTACCACAGTGAAAAGAGCCTCAGACACACCACCACCTCCTCAGC
AGTGGGGCGCCCCCAGTGCCCACTCCTTCCCCACCTCCTTTTACTAAGGGTGTGGTT
ACAGACAGCAAAGTCACATCAGCTTTCCAGATGACGTCAAATAGAGTGGTCACCAT
ATATGAATCTTCAAGGCACAATACAGATCTGCAGCAACCCTCAGCAGAGGCTAGCC
CCAATCCTGAGATCATAACTGGAACCACTGACTCTCCCTCTAATCTGTTTCCATCCAC
TTCTGTGCCAGCACTAAGGGTAGATAAACCACAGAATTCTAAATGGAAGCCCTCTCC
CTGGCCAGAACACAAATATCAGCTCAAGTCATACTCCGAAACCATTGAGAAGGGCA
AAAGGCCAGCAGTAAGCATGTCCCCCACCTCAGCCTTCCAGAGGCCAGCACTCAT
GCCTCACACTGGAATACACAGAAGCATGCAGAAAAGAGTGTTTTTGATAAGAAACC
TGGTCAAAACCCAACCTTCCAAACATCTGCCTTACGTCTCTCTACCTAAGACTCTATTG
AAAAAGCCAAGAATAATTGGAGGAAAGGCTGCAAGCTTTACAGTTCCAGCTAATTC
AGACGTTTTTCTTCCCTTGTGAGGCTGTTGGAGACCCACTGCCCATCATCCACTGGACC
AGAGTTTCATCAGGANTTGAAATATCCCAAGGGACACAGAAAAGCCGGTTCCACGT
GCTTCCCAATGGCACCTTGTCCATCCAGAGGGTCAGTATTCAGGACCGTGGACAGTA
CCTGTGCTCTGCATTTAATCCACTGGGCGTAGACCATTTTTCATGTCTCTTTGTCTGTG
GTTTTTTACCCGGCAAGGATTTTGGACAGACATGTCAAGGAGATCACAGTTCACTTT
GGAAGTACTGTGGAATAAAGTGCAGAGTGGAGGGTATGCCGAGGCCTACGGTTTC
CTGGATACTTGCAAACCAAACGGTGGTCTCAGAAACGGCCAAGGGAAGCAGAAAGG
TCTGGGTAACACCTGATGGAACATTGATCATCTATAATCTGAGTCTTTATGATCGTG
GTTTTTACAAGTGTGTGGCCAGCAACCCATCTGGCCAGGATTCAGTGTGGTTAAGA
TACAAGTCATCACAGCTCCCCCTGTCATTATAGAGCAAAAGAGGCAAGCCATCGTTG
GGGTTTTAGGTGGAAGTTTGAAACTGCCCTGCACTGCAAAAGGAACTCCCCAGCCTA
GTGTTCACTGGGTCCCTTATGATGGGACTGAACTAAAACCATTGCAGTTGACTCATT
CCAGATTTTTCTTGTATCCAAATGGAACCTCTGTATATAAGAAGCATCGCTCCTTCAGT
GAGGGGCACTTATGAGTGCATTGCCACCAGCTCCTCAGGCTCAGAGAGAAGGGTAG
TGATTCTTACTGTGGAAGAGGGAGAGACAATCCCCAGGATAGAAACTGCCTCTCAG
AAATGGACTGAGGTGAATTTGGGTGAGAAATTACTACTGAACTGCTCAGCTACTGG
GGATCCAAAGCCTAGAATAATCTGGAGGCTGCCATCCAAGGCTGTCATCGACCAGT
GGCACAGAATGGGCAGCCGAATCCACGTCTACCCAAATGGATCCTTGGTGGTTGGG
TCAGTGACGGAAAAAGACGCTGGTGACTACTTATGTGTGGCAAGAAACAAAATGGG
AGATGACCTAGTCCTGATGCATGTCCGCCTGAGATTGACACCTGCCAAAATTGAACA
GAAGCAGTATTTTAAGAAGCAAGTGCTCCATGGGAAAGATTTCCAAGTTGACTGCA

AGGCCTCTGGCTCCCCTGTGCCTGAGGTATCCTGGAGTTTGCCTGATGGGACAGTGC
TCAACAATGTAGCCCAAGCTGATGACAGTGGCTATAGGACCAAGAGGTACACCCTT
TTCCACAATGGAACCTTGTATTTCAACAACGTTGGGATGGCAGAGGAAGGAGATTAT
ATCTGCTCTGCCCAGAACACCTTAGGGAAAGATGAGATGAAAGTCCACCTAACAGT
TCTAACAGCCATCCCACGGATAAGGCAAAGCTACAAGACCACCATGAGGCTCAGGG
CTGGAGAAACAGCTGTCCTTGACTGCGAGGTCACTGGGGAACCGAAGCCCAATGTA
TTTTGGTTGCTGCCTTCCAACAATGTCATTTCAATTCTCCAATGACAGGTTACATTTTC
ATGCCAATAGAACTTTGTCCATCCATAAAGTGAAACCACTTGACTCTGGGGACTATG
TGTGCGTAGCTCAGAATCCTAGTGGGGATGACACTAAGACATACAACTGGACATT
GTCTCTAAACCTCCATTAATCAATGGCCTGTATGCAAACAAGACTGTTATTAAAGCC
ACAGCCATTTCGGCACTCCAAAAAATACTTTGACTGCAGAGCAGATGGGATCCCATCT
TCCCAGGTCACGTGGATTATGCCAGGCAATATTTTCTCCCAGCTCCATACTTTGGA
AGCAGAGTCACGGTCCATCCAAATGGAACCTTGAGATGAGGAACATCCGGCTTTC
TGACTCTGCGGACTTCACCTGTGTGGTTCGGAGCGAGGGAGGAGAGAGTGTGTTGGT
AGTGCAGTTAGAAGTCCTAGAAATGCTGAGAAGACCAACATTCAGAAACCCATTCA
ACGAAAAAGTCATCGCCCAAGCTGGCAAGCCCGTAGCACTGAACTGCTCTGTGGAT
GGGAACCCCCACCTGAAATTACCTGGATCTTACCTGACGGCACACAGTTTGCTAAC
AGACCACACAATTCCCCGTATCTGATGGCAGGCAATGGCTCTCTCATCCTTTACAAA
GCAACTCGGAACAAGTCAGGGAAGTATCGCTGTGCAGCCAGGAATAAGGTTGGCTA
CATCGAGAAACTCATCCTGTTAGAGATTGGGCAGAAGCCAGTCATTCTGACATACGA
ACCAGGGATGGTGAAGAGCGTCAGTGGGGAACCGTTATCACTGCATTGTGTGTCTG
ATGGGATCCCCAAGCCAAATGTCAAGTGGACTACACCGGGTGGCCATGTAATCGAC
AGGCCTCAAGTGGATGGAAAATACATACTGCATGAAAATGGCACGCTGGTCATCAA
AGCAACAACAGCTCACGACCAAGGAAATTATATCTGTAGGGCTCAAAACAGTGTTG
GCCAGGCAGTTATTAGCGTGTGTCAGTGATGGTTGTGGCCTACCCTCCCCGAATCATAA
ACTACCTACCCAGGAACATGCTCAGGAGGACAGGGGAAGCCATGCAGCTCCACTGT
GTGGCCTTGGGAATCCCCAAGCCAAAAGTCACCTGGGAGACGCCAAGACACTCCCT
GCTCTCAAAGCAACAGCAAGAAAACCCCATAGAAGTGAGATGCTTCACCCACAAG
GTACGCTGGTCATTCAGAATCTCCAAACCTCGGATTCCGGAGTCTATAAGTGCAGAG
CTCAGAACCTACTTGGGACTGATTACGCAACAACCTTACATCCAGGTACTCTGACAGG
AAGGGGGAGACTAAAATTCAACAGAAGTCCACATCCACAGGGTTTATTTTTTGGA
GAAGTTTAATCAAAGGCAGCCATAGGCATGTAAATGAGTCTGAATACATTTACAGT
ATTAAATTTACAATGGACATGCGATGAGACTTGTAATGAAAGCATTGTGAACTGA
AACCGAGTCTCTGTGGATCTCAAAGCAAACCTTAACTTAAGGCACTTTGATTTTGC
CAACAAATAATAACAAACATTAAGAGAAAAAAATGATCCACTACGAAATAACAAAC
GGCTAATGCACCTGAATTCTCAGTAAAAAGACCTTTCTCTCGCTAACAGTTGCCAGC
TGCCTCGTGTCTGTTTCTACCAATGTCACAAACATCGCACACAGGGTGAATGGAGT
CAACGGGAAAGATTAAGTTTGCAGTCTGTGTAATCTCAATGTACAAATATTCTGTC
NCTGGTTTATAAACATTTTGATAAAACCGAAAAA
AAAA (SEQ ID NO: 1)

Figure 2

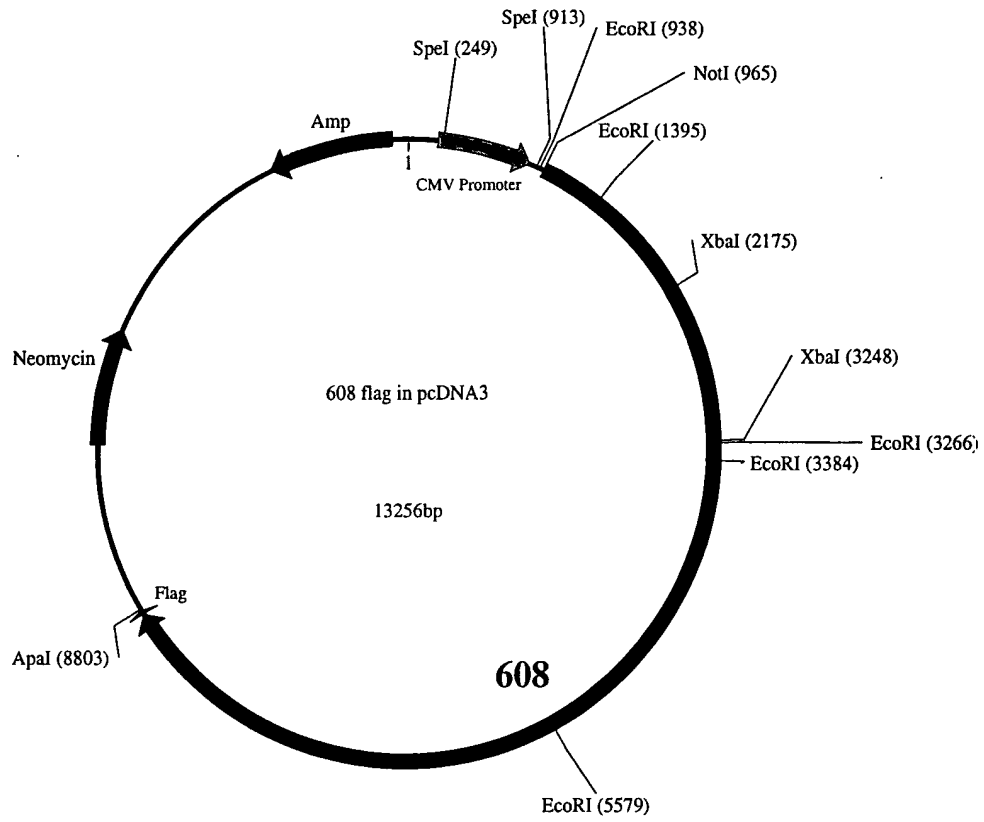


Figure 3

MQVRGREVSGLLISLTA VCLVVT PGSRACPRRCACYVPTEVHCTFRYLTSIPDGIPANVE
 RINLGYNLTRLTENDFDGLSKLELLMLHSNGIHRVSDKTFSGLSLQVLKMSYNKVQII
 RKDTFYGLGSLVRLHLDHNNIEFINPEAFYGLTSLRLVHLEGNRLTKLHPDTFVSLSYLQI
 FKTSFIKYLFLSDNFLTSLPKEMVSYMPNLESYLHGNPWTCDCHLKWLSEWMQGNPDI
 IKCKKDRSSSSPQQCPLCMNPRISKGRPFAMVPSGAFLCTKPTIDPSLKSLSLVTQEDNGS
 ASTSPQDFIEPFGSLSLNMTXXSGNKADMVCSIQKPSRTSPTAFTEENDYIMLNASFSTNL
 VCSVDYNHIQPVWQLLALYSDSPLILERKPQLTETPSLSSRYKQVALRPEDIFTSIEADVR
 ADPFWFQQEKIVLQLNRTATTLSTLQIQFSTDAQIALPRAEMRAERLKWTMILMMNPNK
 LERTVLVGGTIALSCPGKGDPSPHLEWLLADGSKVRAPYVSEDGRILIDKNGKLELQMA
 DSFDAGLYHCISTNDADADVLT YRITVVEPYGESTHDSGVQHTVVTGETLDLPCSTGV
 PDASISWILPGNTVFSQPSRDRQILNNGTLRLQVTPKDQGHYQCV AANPSGADFSSFKV
 SVQKKGQRMVEHDREAGGSGLGEPNSSVSLKQPASLKL SASALTGSEAGKQVSGVHRK
 NKHRDLIHRRRGDSTLRRFREHRRQLPLSARRIDPQRWAALLEKAKKNSVPKKQENTTV
 KPVPLAVPLVELTDEEKDASGMIPPDEEFMVLKTKASGVPGRSPTADSGPVNHGFMTSI
 ASGTEVSTVNPQTLQSEHLPDFKLFSVTNGTAVTKSMNPSIASKIEDTTNQNPIIFPSVAE
 IRDSAQAGRASSQSAHPVTGGNMATYGHNTYSSFTSKASTVLQPINPTESYGPQIPITGV
 SRPSSSDISSHTTADPSFSSHPSGSHTTASSLFHIPRNNNTGNFPLSRHLGRERTIWSRGRV
 KNPHRTPVLRHRHRTVRPAIKGPANKNVSQVPATEYPGMCHTCPSAEGTLVATAALS
 VPSSSHSALPKTNNVGVI AEESTTVVKKPLLLFKDKQNV D I E I T T T T K Y S G G E S N H V I P T E
 ASMTSAPTSVSLGKSPVDNSGHL SMPGTIQTGKDSVETTP L S P L S T P S I P T S T K F S K R K T P
 LHQIFVNNQKKEGMLKNPYQFGLQKNPA AKLPKIAPLLPTGQSSPSDSTLLTSPPPALST
 TMAATQNKGTEVVSGARSL SAGKKQPFTNSSPVLPSTISKRSNTLNFLSTETPTVTSPTAT
 ASVIMSETQRTSRKEAKDQIKGPRKNRNNANTTPRQVSGYSAYSALT TADTPLAFSHSP
 RQDDGGNVSAVA YHSTTSLLAITELFEKYTQTLGNTTALETTLLSKSQUESTTVKRASDTP
 PLLSSGAPPVPTSPPPFTKGVVTD SKVTS AFQMTSNRVVTIYESSRHNTDLQQPSAEAS
 PNPEIITGTTDSPSNLFPSTSVPALRVDKPQNSKWKPSWPPEHKYQLKSYSETIEKGKRP
 VSMSPHLSLPEASTHASHWNTQKHA EKS VFDKKPGQNPTSKHLPYVSLPKTLLKKPRIIG
 GKAASFTVPANS DVFLPCEAVGDPLPIIHWTRVSSGXEISQGTQKSRFHVLPNGTLSIQRV
 SIQDRGQYLCSAFNPLGVDHFHVSLSVVFYPARILDRHVKEITVHFGSTVELKCRVEGMP
 RPTVSWILANQTVVSETAKGSRKVWVTPDGT L I I Y N L S L Y D R G F Y K C V A S N P S G Q D S L L
 VKIQVITAPPVIEQKRQAIVGVLGGS L K L P C T A K G T P Q P S V H W V L Y D G T E L K P L Q L T H S
 RFFLYPNGTLYIRSIAPSVRGTYECIATSSSGSERRVILTVEEGETIPRIETASQKWTEVN
 LGEKLLLNC SATGDPKPRIIWRLPSKAVIDQWHRMGSRIHVYPNGSLVVGSVTEKDAGD
 YLCVARNKMGDDLVL MHVRLRLTPAKIEQKQYFKKQVLHGKDFQVDCKASGSPVPEV
 SWSLPDGTVLNNVAQADDSGYRTKRYTLFHNGTLYFN NVGMAEEGDYICSAQNTLGK
 DEMKVHLTVLTAIPRIRQSYKTTMRLRAGETA VLDCEVTGEPKPNVFWLLPSNNVISFS
 NDRFTFHANRTLSIHKVKPLDSGDYVCVAQNPSGDDTKTYKLDIVSKPPLINGLYANKT
 VIKATAIRHSKKYFDCRADGIPSSQVTWIMPGNIFLPAPYFGSRVTVHPNGTLEMRNIRLS

DSADFTCVVRSEGGESVLVVQLEVLEMLRRPTFRNPFNEKVIAQAGKPVALNCSVDGNP
PPEITWILPDGTQFANRPHNSPYLMAGNGSLILYKATRNLKSGKYRCAARNKVGYIEKLIL
LEIGQKPVILTYEPGMVKSVSGEPLSLHCVSDGIPKPNVKWTTPGGHVIDRPQVDGKYIL
HENGTLVIKATTAHDQGNYICRAQNSVGQAVISVSVMVVA YPPRIINYLPNMLRRTGE
AMQLHCVALGIPKPKVTWETPRHSLLSKATARKPHRSEMLHPQGTLVIQNLQTSDSGVY
KCRAQNLLGTDYATTYIQVL (SEQ ID NO: 2)

U.S. Patent & Trademark Office

Figure 4

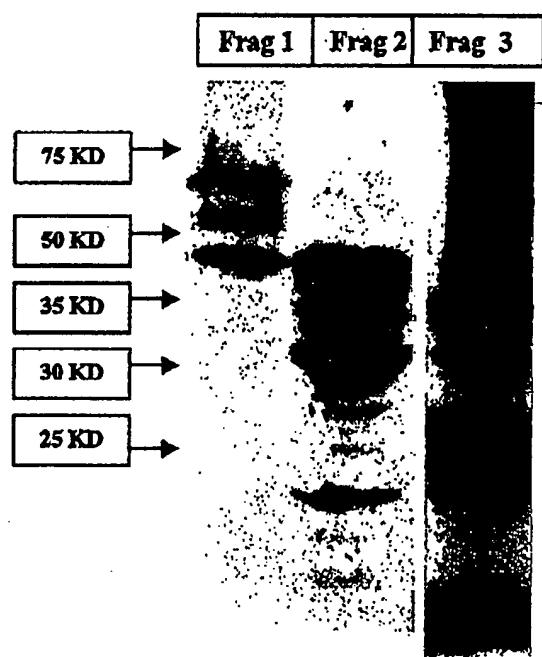


Figure 5

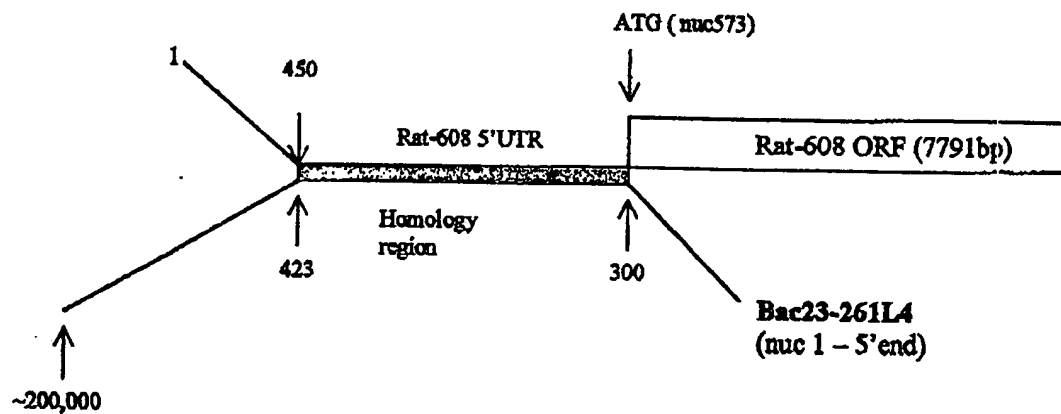
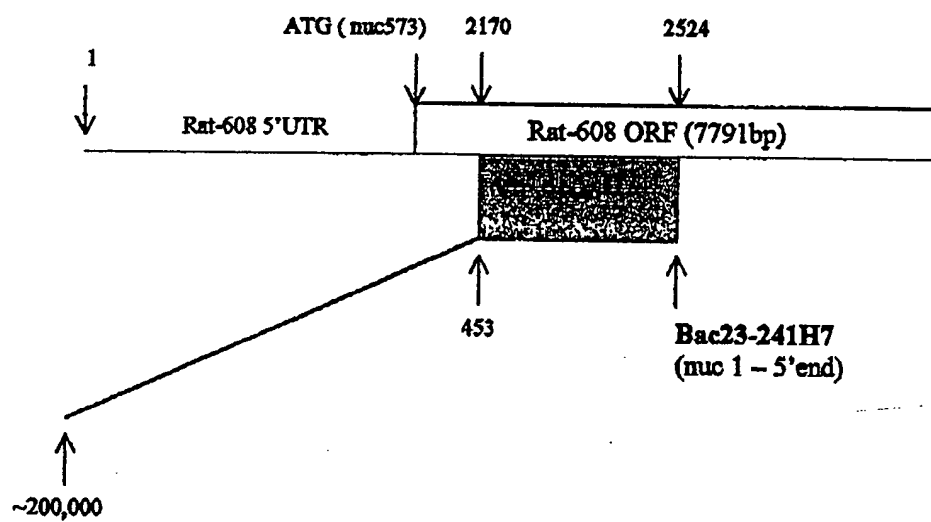


Figure 6



Nbc 1

↓ Nuc 390

▲ Nec 2817

AAAAAAAAAAAAAAAAAAGGAAGGTCAGACACCTCACTTCACTGCTATCTCAACTT
GCAAACAGAAGGGGAGTCACAAACCCAGGACAAACCACAGTGATTGAAGCGTCTTTGAA
TGTTATTGCTGTTGTTGTTACCAACATCATTAGCATATATTCATTGTGAAAACCTTACGGGGT
CTATGCATGTTTTTTTATTCAAGTATATCACATGCTGTCAGCATATTTGGCACCCTACCA

▼ Exon 1 (Nuc 3661)

↑ Nuc 4203

TGTATAAATACAGATTTCAGGCTTAAGTGTACGGGAACTCATGTGGTATTCAATTAACGGCTCT
 CTCTCTTATAACTAAGCTTAAGGTGCATATAGTCTCTCTGTTTCCAGCTACCTTGTAC
 CATCTTTGTTTATCTAATAATAGCAAGCTCATCTGCTTTTAATCATCACGCAGAGAGTATT
 CAAAAATATTCAGTATGTAACAGTGCAGCTGTAGGCATAGGAAGTAATCATTAGTAAATC
 TTAATTTGGGTAAACTCATTCATACAGCTOCAGGTGGGAGGATCACTGAGGCTTGC3C
 CAGTGC0GGGTAAAGATATTTTCTAACAGAGAAGCAGAATTCCTCTTGCCCATGCTCC
 CCATCACTGTGTCAGTAAGCAGAGGGGTGTTTCCAAGCAGAGAAAGAGCAGACAGTGT
 TGCTGCAAAAGTCAGAGACTCAGCCCTCCAGCTGGTCAGTTTACTGTCTCCCGGTCAAT
 AGTTGGCTCTGAAAAGGCCCATGTGTCTTATTGGCAAGGACTTGCAGACATGCTAGAAA
 GAAATTTGACCTTTTCTAGATGGGTATTACAGCTGTAAAAGTATTTTGGAAAGTTAAG
 CCAATAAATAAACAACATATTAATAATCAATGTACAAAATTTGATCATATAAAGAA
 GTACATTCATAAATGCAATGTGAAAAATATATAATAATTTTATCTATTACTGGTGCAAA
 TTTTCTAAATGTCAGTACTATTTTATATTATTAATAAAATATTTTAAATGTATATAA
 AGTGTA AAAAGGCTCTTGGTCAAACAAGAGATTAAATTTACAAACTTTAATTGTCCCAT
 AACATTATTATGATCTCTAATGACAGGGATCCTGCTTTTCATTGGGAAATGAGAAGCTATG
 AAGATATGTTTACAATAATAAGCCATTTAGTGATAAAGTCCAAATGGGAAGCTAGCACAC
 ACTGGTTTATAAAGAGAACAGTTTCTCTGAGTCTATGCAAGTTTACACTCTAGGGATAAG
 AGTTCCTCTTCTCCAGATTTCACAGTACTTTGTGTCATCATTTTATCTCTTGATGATGAG
 CATTATAAGTGGAAATAAGATAGGATCTCAAAGGAATGTCAATTTGGATGCCCTGAACAAT
 CTTTCAGGCTTTCTTTTCAGTTCACTAGTCTATTCATTTATTGGATAATTGGGGGATGGTGT
 TAATTTTTTTGCAGTCTTATGGAATTCCAAAAACAACAAAAACAACAACAAAAA
 AACTCTGAAACTAGAACTACCAATCCATTACTGGGTATGTAAACAAAGAGAAATCTGCACA
 GAATTTATTGCTACATTGTTTCATTATTCACGACAGCCAAGAATGTGGAACCAACTTACGTA
 GCGGTCAAAATATGAAAGGATAAAGAAAAATGTGGAATGTGTACAACAGAGTCCCATGTG
 GCCATAAAGATGTGAAATCATAGCATATGCAAGAAATGGATGTCAACTGGAATCAATTT
 GGCTAATCAAAACAAGACAGAGCTCAAAAAGGAAACACCGGTGAGCTTCTCTGACAAACA
 GAAGCTAGATTTACACTTGTACGTGGCATGTGTGTTTGAATTTTATTAGTTATACACT
 ATTCTAATCTGTGAGTGTGTATAAAGGCATGCATGTAAAGCAAAAACAAGCTAGCTGGG
 TGGTATAGGAGAGAAAAGCAATGAGAGGAGTTAATAAGAACGAAGCATAGTAACATAGGTG
 CCAGGATGAAATGCATTAATTTGTATGCTAACTAAACCAACAGACAGGAGGCCACAGTTCA
 AACAGGGTGAAATCCAGCACAGAGAAGGGAAGTAGACACAAGTTTCGCCACTAAC
 CAAAGAAGCCATTTCAGTTGCTGCTGCTGCTGGGAGGGGCTTCCAGTTTCTCCAGTCTGAC
 ACTGTGTATAACAACCAAGTTGACAATACAAAGTTGGCATGATGGATGGTTTTTGTGCTATT
 TTTCATTTTCTTCTTACGTGTTTGTGTGTGTGGTGGTGTGTGTGGTGGTGGCTGTGGTTTTT
 ATTGTGTTCTTTTGAGAGAGAGAAAGGAACATGAAATTTGGGTGGGTAGGAAGCTGGAACCG
 ATCTGGAAGAAAGTTGGGGAAAGAAAAATTTGATTTGGAGCATATTTAAACAAACAAACA
 AACAAACAAAGGTTCAATTTGCCACAAAAGGTGTGAATTAATTAACAGTTACGACT
 CTTAAAGAAAATATTCCCAATTATTCCAGAGTTGCTATGTATGCTGTGCTTAGGACTTTG

▼ Exon 2 (Nuc 6559)

CTTGAAC TGGCCCTATAACTCTGGTGTGGTGTCTTTTCAGGATGCAGAAAGAGAGGCAGGGCAA
 GTCAGCTGCTGCTGATCTCCCTCACTGCCATCTGCTGGTGGTCACCCCTGGGAGCAGGGTC
 TGTCTCGCGATGTGCTGCTATGTGCCACAGAGGTGCACTGTACATTTGGGGACCTGACC
 TCCATCCAGACGGGCATCCAGCCAATGTGGAACGAGTCAATTTAGGGTGTGT

↑ Nuc 6755

GGACCTTGCTGATCTCCTTCTCAGAGAGGGACCACTGATTTTCTGCTACTTTGCCCCCC
 AAACACCTGTGATTACTTTTAATAGTTTTCTTCTAAAATGGGTTTCATACAAACCTTATAITG
 TGGAGACAATGAACATTTTATCCCAATAGTCTTTTACTAGAACTTGAAGCCCTCTTAGT
 GTTTGGGAGCCTCATAATTATGGGGCAGCTTTATTCTGAATGAATTTTAAATGAAAAAGAT
 ACAGTTTCTGTAAACAATCATTATGATACCAAGGAAGAGGAATTGTCAATTGAATATTTTAA
 AAAAGCATTTCTTTTGAATTTATAAATACCCATTACAAAATGGCTTACTTAAAATACTTG
 CCTTACTAAATCTGACAAATTATGGTGATATTTTGAAGGTTTATGAAAAATTTGTTTATGT
 ATAAATGCACAAGAAATGGGATATGCCATCACCTATGTCCATTAGTGAGCATGTACAGT
 ATGCCAAACACTATTGTTTCAGTTTGGAGGAAGTAATGGGGGTGGGGGAGCAACAAGGGT
 TATAACCGTATACCCAGTGCTTGGGAAGCGATTGCAAAACAGTAAAGACTGACATTGTGTT
 CTCCCTATGAGGGAGGGGCTTGGGCTGAGCACTTTGCAATGAGCATTGTCTCATTGTGCT
 GGCAGGTTTTATGATAACTTGACCCAAGCTAGAGTCACTGGAGAGGAAGGAACCTCAACT
 GAGAACATGCCTGAAGAAGATCAGATTATAGGCAGGCTGTGGGGCATTTTCTTAATTAG
 TGATTCATGGGGCAGGGCCAGTCCATTGTTCTGTGTACCATTCTCAGGCATTTAAAAA
 AAAAAAAAAACAGGCTGAGCAAGTGTCAAGGAGCAAGTCAGTGAGCAGCAGCCCTAATGA
 TCTCTGCATCAGCTCCTGCCCTCCAGGTTCTACCTATTTGAGTTCCTGTCTAGCTCCCTA
 CAGTGATGAACAATGATGTGGAAATATAAGCCAAATAAATCCTTTCTTCCCAACTTGCTG
 TTGGTCATGATGTTTCATCAGTGATAATAGTCTCATGAAGATGCTGGTGTTTATAACA
 CCTTTGGACTAAATTCTGTTATCTATAGCTGAGGAAAAATGGAGCATAGAAAGTCTCCAGA
 CTACACCAGAGTGTAACTGGGCTGAGCTTAGAATCACCCACGTGCACTCCACTGCC
 GGGGCTTCTTAACCGGAACACAGTTGTAAAAGGGAATTTTCTGTTTGTTCATTTTGACA

Exon 3 (Nuc 8089) ▼

TGTGGACTTTAATTGACGATTTCATCTGAAGCTGAAAATGATTTTTTTCCAGGTATAACAGC
 CTCACTAGATTGACAGAAAATGACTTTTCTGGCTGAGCAGACTGGAGTTACTCATGCTGCACA
 GCAATGGCATTTCACAGAGTCAGTGACAAGACCTTCTGGGCTTGCAGTCTTGCAG

Nuc 8218 ↑

GTGAGATAGGTAGAGGGTGATGGAGGCTGAGAAGAGAGGGTGCAACTGTGGGTATACCC
 AAAAGCTGCTGATTCCCGTGGGAGACATTCTATAAGCATTCTATAAACTAGAGGCAGATA
 TCAAGGAAGGATTTCAATTGTAATGCAATTTTATGAGAAAATTTGAATATTAAGAAAATG
 CTGGGGAAAATGCTTACACAATTGCGAGGACCTAATTTAGGATCTCCAATAGCCACATAA
 AAAGCACAGCATGGCGGCAGACACCTGCAATTCCTGTCCCTGGAAGCACCTGTTCAGAAT
 CCCAGAGACTCATTGGCCAAACACTCTATTCAATCAATGAAGTCCATATTCACTGACAAA
 ACTTGACTCAGAACTAATGTGGAAAGCATCAGGAAGACAGCCAACTCTGCTCTACT
 CATGCAATGAATAAGGATCCAGAGAGAAAGGAAAGAAAAGGAAGGAAGGAAGGAGG
 AAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAGGGAGGAAAGGAGGGAGGG
 AAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGG
 GGAGAAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGG
 TTTTATCCATAAAAGGTCAATTTCCACCTGTTTGTCTGGAAGTAGAGTGGGATCCCTTATAT
 AAGGGCAGTCTTTAACATAGTAGCATTTTATAAACCATTACAAATTTGAGTTTTCTCTAC
 TTTTATCTCTACCATCTTCAAACCTGAAACTACAATTATTOCCACAAATGAAGAAAATGC
 TGTAAAGATTTTACACACCCGAAGTGGGAACTTAAGGATTAGACAAGTCTAACAATGAG
 AATGGGGAGAACAAAAAGAGACTGCACAGGGAGCCCTTTCTCTGCTTATAATCTTGACAC
 TTGAGAAAGCTAATTGACGCTGCATGACTACTCAACTCTTTAAGCAACAATGCTGTTGTTC
 ATGAAAAGCACAATAAAGTACATATGTCCATAATATTCATCAAAATTTGCATGCAGCAC
 ATAATAGCAATCAAAGCAATAACCCCACTGTTTCACAGAGACTTTAAACATGAAACTGGA
 ACTATGTCTAGTGTGTTGACTTAGGGTACATAGTATGCTGTGTCTGTATGTACCAATGTG
 ATTTAGGTCATCAGACAGCATTGGAACATGTATCTTCAGGAGGAATCATTGATGTATCT
 GCATGAAATTTCTCCACCTATGTTTATTCTCTTAGCCAGGTTTTCTCTGATGGAGAAACATT
 GGGTTTGAGGTTTTACTCCAGGTAACATTTAGGGAAAAGCTGTCTATGTTCTCAGTTTGG
 CTTTTATTATGAGGGATGTTGGTATTCCAGAAAATTTCTTTTGAAGAGATTACAATTTA
 GGTCAAAACAGAAAAATATGTAAAAAGTTATTGTTTTATTAGTATTTCATGTTCTTTCTT
 TTTTAAAAATGGTATGCTTAGAACTAATTAAGATTAGATTAGATTAGATTAGAAAATAATC
 AGAGAGGGATTGATGAATGCTAAAGCATCATGAAAAATTCAAAATTTTGTCTCTAATT
 CAGAATCAATTAATTCATATTACTATAAAAGACAGCACGCCAGATGTGTGCCAGCTGAG
 GAGTGGATAAACTGTGTAACGTGAGTGCTATGTAGAAACAGAAAGGAAGTGAAGGGTTGA

Exon 4 ? (Nuc 11286) ↓

Nuc 11680

Nuc 11967 ▲

(SEQ ID NO:3)

Figure 8

Exon/Intron No.	Exon start	Exon end	Exon length	Intron length
1	1	542	542	2356
2	2898	3094	197	1334
3	4428	4557	130	3068
4	7625	8019	394	

Figure 9

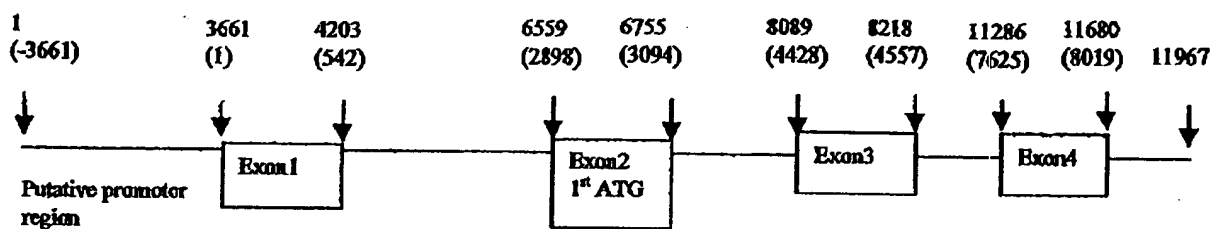


Figure 10

cDNA_rat	1	ogagagacgacagaagggttacggctgcgagaagaacagacagaagggtccag
genomic_hu	1	-----
cDNA_rat	51	aaaaaggaaagtgcctggaggggagtgaggacaaaagcagcgaccaagtga
genomic_hu	1	-----
cDNA_rat	101	atgtcacttcagtgcactgagggcaggcaaaaacggcggaaggattttgt
genomic_hu	1	-----
cDNA_rat	151	gtagcttgggaocctttcatagacactgatgacacgtttacgaaaaatag
genomic_hu	1	-----
cDNA_rat	201	aaatttgaggagaaaacgcctgggccttoggaaaggagtgattgattagta
genomic_hu	1	-----
cDNA_rat	251	cttgaagtttaggtgactttaaggagaactaactaatgtatactattga
genomic_hu	1	-----
cDNA_rat	301	gggaggaggaagagcattacagagtttcagcagcagcaggaaagctttg
genomic_hu	1	-----
cDNA_rat	351	gttaatttggaatggatgatagcattaaaaataacagaagcgctccag
genomic_hu	1	-----
cDNA_rat	401	tototgaagottoagtccccagctgaaagccagaaaaagaotaagccaa
genomic_hu	1	-----
cDNA_rat	451	taagccttttgatccctttggaagcaaagaactttccttcctggggtga
genomic_hu	1	-----
cDNA_rat	501	agactctcctcagaagatttcctgtctctgcctatgttacaagaggaatc
genomic_hu	1	-----
cDNA_rat	551	aaaaccaagacagaagagctcaggatgcaggtgagaggcagggaagtca
genomic_hu	1	-----
cDNA_rat	601	oggottgtgatctocctoactgctgtctgcotggtggtcacccctggga
genomic_hu	1	-----
cDNA_rat	651	gcagggcctgtcctgcgcgtgtgctgtatgtgcccaagaggtgca
genomic_hu	1	-----
cDNA_rat	701	tgtacatttoggtaacctgacctccatccagatggcatcccgccaatgt
genomic_hu	1	-----
cDNA_rat	751	ggaacgaataaatttaggatataacagccttactagattgacagaaaac
genomic_hu	1	-----
cDNA_rat	801	actttgatggcctgagcaaaactggagttactcatgctgcacagtaatgg
genomic_hu	1	-----
cDNA_rat	851	attcacagagtcagtgacaagaccttctcgggcttgcaagtcttgcaggt
genomic_hu	1	-----
cDNA_rat	901	ottaaaaatgagotataaocaaagtocaaatoattoggaaggatactttct
genomic_hu	1	-----

cDNA_rat	951	acggactcgggagcttgggtccggttgacacctggatcacacaacacattgaa
genomic_hu	1	-----
cDNA_rat	1001	ttcatcaaccctgaggccttttatggacttacctcgctccgcttggtaca
genomic_hu	1	-----
cDNA_rat	1051	tttagaaggaaacccggctcacaaagctccatccagacacatttgtotcat
genomic_hu	1	-----
cDNA_rat	1101	taagctatctccagatatttaaaccctctttcattaagtacctgttcttg
genomic_hu	1	-----
cDNA_rat	1151	tctgataacttctctgacctccctccaaaagaaatggtctcctacatgcc
genomic_hu	1	-----
cDNA_rat	1201	aaacotagaaagcctgtattttgcatggaaacccatggacctgtgactgcc
genomic_hu	1	-----
cDNA_rat	1251	atttaaagtggttggtotgagtggtatgcagggaacccagatataataaan
genomic_hu	1	-----
cDNA_rat	1301	tgcaagaagagacagaagotcttccagtcctcagcaatgtccctttgcat
genomic_hu	1	-----
cDNA_rat	1351	gaaccccaggatctctaaaggcagaccctttgctatggtacocatgtggag
genomic_hu	1	-----
cDNA_rat	1401	ctttcctatgtacaaagccaaccattgatccatcactgaagtcacagagcc
genomic_hu	1	-----
cDNA_rat	1451	ctggttactcaggaggacaatggatctgctccacctcacctcaagatttt
genomic_hu	1	-----
cDNA_rat	1501	catagaaaccccttggctccttgtctttgaacatgacanannntnctggaa
genomic_hu	1	-----
cDNA_rat	1551	ataaggccgacatggtctgtagtatccaaaagccatcaaggacatcacca
genomic_hu	1	-----
cDNA_rat	1601	actgcattcactgaagaaaatgactacatcatgtaaatgcgtcatttttc
genomic_hu	1	-----
cDNA_rat	1651	cacaaatottgtgtgcagtgtagattataatcacatccagccagtggtggc
genomic_hu	1	-----
cDNA_rat	1701	aaottctggctttatacagtgactctcctctgatactagaaaggaagccc
genomic_hu	1	-----
cDNA_rat	1751	cagcttaccgagactccttcactgtcttctagatataaacaggtggctot
genomic_hu	1	-----
cDNA_rat	1801	taggcctgaagacatttttaccagcatagaggctgatgtcagagcagacc
genomic_hu	1	-----
cDNA_rat	1851	ctttttggttccaacaagaaaaattgtcttgcagctgaacagaactgcc
genomic_hu	1	-----
cDNA_rat	1901	accacacttagoacattacagatocagttttccactgatgtctcaaatccg
genomic_hu	1	-----
cDNA_rat	1951	tttaccaagggcggagatgagagcggagagactcaaatggacocatgatoc
genomic_hu	1	-----

cDNA_rat genomic_hu	2001 1	tgatgatgaacaatccccaaactggaaagcaactgttoctgggttgccggcact -----
cDNA_rat genomic_hu	2051 1	attgccotgagctgtccagggoaaaggogaccccttcacctcacttgggaatg -----
cDNA_rat genomic_hu	2101 1	gottotagctgatgggagtaaagttagagcccccttacggttagcagaggatg -----
cDNA_rat genomic_hu	2151 1	ggcgaatcctaataagacaaaaatgggaagttggaactgcagatgggtgac -----
cDNA_rat genomic_hu	2201 1	agctttgatgcaggtotttaccactgcataagcaccaatgatgcagatgc -----
cDNA_rat genomic_hu	2251 1	ggatgtttcacatacaggataactgtggttagagccctatggagaaagca -----
cDNA_rat genomic_hu	2301 1	cacatgacagtggagtcacgacacagtggttacgggtgagacgctcgao -----
cDNA_rat genomic_hu	2351 1	cttccatgcctttccacgggtgttcagatgottotattagctggattct -----
cDNA_rat genomic_hu	2401 1	tccagggaaacactgtgttctctcagcoactcaagagacaggcaaattctta -----
cDNA_rat genomic_hu	2451 1	acaatgggaccttaagaatattacaggttacgcacaaaagatcaaggatcat -----
cDNA_rat genomic_hu	2501 1	taccaatgtgtgggtgccaaacooatoaggggcccgaactttocagtttta -----
cDNA_rat genomic_hu	2551 1	agtttcagttcaaaaagaaaggccaaaggatgggttgagcatgcaggagg -----
cDNA_rat genomic_hu	2601 1	caggtggatctggacttggagaaocaaactccagtggtttcccttaagcag -----
cDNA_rat genomic_hu	2651 1	ccagcatctttgaaactctctgoatcagctttgacagggtcagagggtgg -----
cDNA_rat genomic_hu	2701 1	aaaacaagtctccggtgtacataggaagaacaaacatagagacttantac -----
cDNA_rat genomic_hu	2751 1	atcggcggcggtggggattccacgctccggcgattcaggggagcataggagg -----
cDNA_rat genomic_hu	2801 1	cagctccctctctctgctcggagaattgaccgcgaacgctgggcagcact -----
cDNA_rat genomic_hu	2851 1	tctagaaaaagccaaaaagaattctgtgccaaaaaagcaagaaaatacca -----
cDNA_rat genomic_hu	2901 1	cagtaaaagccagtgccactgggtgttccctcgtggaactcactgaogag -----
cDNA_rat genomic_hu	2951 1	gaaaaggatgcctctggcatgattcctccagatgaagaattcatggttct -----

cDNA_rat	3001	gaaaaactaagggtcttctgggtgtccacaggaagggtcaccaactgctgactctg
genomic_hu	1	-----
cDNA_rat	3051	gaccagtaaatcatggtgtttatgaacgagtatagcttctggcacagaagtc
genomic_hu	1	-----
cDNA_rat	3101	tcaactgtgaatocacaaacactacaatctgagcaccttootgatttcna
genomic_hu	1	-----
cDNA_rat	3151	attatttagtgtaacaaacgggtacagctgtgacaaagagtatgaacccat
genomic_hu	1	-----
cDNA_rat	3201	coatagcaagcaaaatagaagatacaacocaaocaaaacccaatcattatc
genomic_hu	1	-----
cDNA_rat	3251	tttccatcagtagctgaaattcgagattctgtctcaggcagggaagagcato
genomic_hu	1	-----
cDNA_rat	3301	ttcccaaagtgcaacacctgtaacagggggaaacatggctacctatggcc
genomic_hu	1	-----
cDNA_rat	3351	atacoaacacatatagtagctttaccagcaaaagccagtagctcttgca
genomic_hu	1	-----
cDNA_rat	3401	ccaataaatccaacagaaagttatggacctcgatacctattacaggagt
genomic_hu	1	-----
cDNA_rat	3451	cagcagacotagcagtagtgacatctcttctcacactactgcagacccta
genomic_hu	1	-----
cDNA_rat	3501	gcttctcagtgacaccttcaggttcacacaccactgcctogtctttattt
genomic_hu	1	-----
cDNA_rat	3551	caatttcctagaaacaacaatacaggttaacttccctgtccaggcactt
genomic_hu	1	-----
cDNA_rat	3601	gggaagagagaggacaatttggagcagagggagaggttaaaaacccacata
genomic_hu	1	-----
cDNA_rat	3651	gaacccacagttotccgacggcatagacacaggactgtgaggccagcanta
genomic_hu	1	-----
cDNA_rat	3701	aagggaacctgtacacaaaatgtgagccaaagttccagccacagagtagcc
genomic_hu	1	-----
cDNA_rat	3751	tgggatgtgccacacatgtccttccgagaggggtcacagtggtactg
genomic_hu	1	-----
cDNA_rat	3801	cagcaactgtcagttccaagttcatccacagtgccctccccaaaactnat
genomic_hu	1	-----
cDNA_rat	3851	aatgttggggtcatagcagaagagtctaccactgtggtcaagaacccaot
genomic_hu	1	-----
cDNA_rat	3901	gttactattttaaggacaaacaaaatgtagatattgagataataacaacca
genomic_hu	1	-----
cDNA_rat	3951	ctacaaaatatccggaggggaaagtaaccaogtgattoctacggaagca
genomic_hu	1	-----
cDNA_rat	4001	agcatgaacttctgtctcaacatctgtatccctggggaaatctcctgtaga
genomic_hu	1	-----
cDNA_rat	4051	caatagtggtcacctgagcatgcctgggaccatccaaactgggaaagatt
genomic_hu	1	-----
cDNA_rat	4101	cagtggaacaaacaccacttcccagccctcagcacacctcaataaoa
genomic_hu	1	-----

cDNA_rat genomic_hu	4151 1	acaagcacaaaattctcaaaaggagaaaactccottgcacacagatctttgt -----
cDNA_rat genomic_hu	4201 1	aaataaccagaagaaggagggtgtttaagaatccatatcaattoggtt -----
cDNA_rat genomic_hu	4251 1	tacaaaagaaccagcgcaaaaggttccaaaatagctcctctttttacc -----
cDNA_rat genomic_hu	4301 1	acaggtcagagttcccccctcagattctacaactctcttgacaagtcgoc -----
cDNA_rat genomic_hu	4351 1	accagctctgtctacaacaatggotgccactcagaacaagggcactgaag -----
cDNA_rat genomic_hu	4401 1	tagtatacgggtgccagaagtctctcagcaggaagaagcagcccttoaco -----
cDNA_rat genomic_hu	4451 1	aaactcctotccagtgttctctagcaccataagoaagagatctaatacatt -----
cDNA_rat genomic_hu	4501 1	aaacttcttgtcaacggaaacccccacagtgacaagtcctactgtactg -----
cDNA_rat genomic_hu	4551 1	catctgtcattatgtctgaaacccaacgaacaagatccaaagaagcaaaa -----
cDNA_rat genomic_hu	4601 1	gacaaaataaaggggcctcggaagaacagaaaacaacgcaaacaccaoooo -----
cDNA_rat genomic_hu	4651 1	caggoaggtttctggctatagtgcatactcagctctaacaacagotgata -----
cDNA_rat genomic_hu	4701 1	cccccttggttttcagtcattccccacgacaagatgatggttggaatgta -----
cDNA_rat genomic_hu	4751 1	agtgoagttgcttatcaactcaacaacotctctttetggccataactgaact -----
cDNA_rat genomic_hu	4801 1	gtttgagaagtacacccagactttgggaataacaacagctttggaaacaa -----
cDNA_rat genomic_hu	4851 1	cgttggttgagcaaatcacaggagagtaaccacagtgaagagcctcagac -----
cDNA_rat genomic_hu	4901 1	acaccaccaccactcctcagcagtggtgggggccccagtgcccaactcctc -----
cDNA_rat genomic_hu	4951 1	cccacctccttttactaagggtgtggttacagacagcaaaagtcacatcag -----
cDNA_rat genomic_hu	5001 1	ctttccagatgacgtcaaatagagtggtcaccatatatgaatcttcangg -----
cDNA_rat genomic_hu	5051 1	cacaatacagatctgcagcaaccctcagcagaggttagccccaatcctga -----
cDNA_rat genomic_hu	5101 1	gatcataactggaaccactgactctctcttaactctgtttccatccactt -----
cDNA_rat genomic_hu	5151 1	ctgtgccagcactaagggttagataaaccacagaattctaaatggaagccc -----
cDNA_rat genomic_hu	5201 1	tctccctggccagaacacaaatatcagctcagtcatactccgaaccat -----

cDNA_rat genomic_hu	5251 1	tgagaagggcaaaaggccagcagtaagcatgtccccccaactcagccttc -----
cDNA_rat genomic_hu	5301 1	cagaggocagcactcatgcctcacactggaatacacagaagcatgcagaa -----
cDNA_rat genomic_hu	5351 1	aagagtgtttttgataagaaacctgggtcaaaaocaaottccaacatct -----
cDNA_rat genomic_hu	5401 1	gccttacgtctctctacctaagactotattgaaaaagccaagaataattg -----
cDNA_rat genomic_hu	5451 1	gaggaaaggctgcaagctttacagttccagctaattcagaogtttttctt -----
cDNA_rat genomic_hu	5501 1	ccttgtgaggotgttgaggacccactgcccatcatccactggacacaggt -----
cDNA_rat genomic_hu	5551 1	ttcatcagganttgaaatatcccaagggcacagaaaagccggttcacag -----
cDNA_rat genomic_hu	5601 1	tgottcccaatggcaccttggtccatccagagggtcagtattcaggaccgt -----
cDNA_rat genomic_hu	5651 1	ggacagtacctgtgctctgcatttaatccaotgggcgtagaccattttca -----
cDNA_rat genomic_hu	5701 1	tgtctctttgtctgtggtttttaccggcgaaggattttggacagacatg -----
cDNA_rat genomic_hu	5751 1	tcaaggagatccacagttcactttggaagtactgtggaactsaagtgcaga -----
cDNA_rat genomic_hu	5801 1	gtggagggtatgccgaggcctacggtttctctggatacttgcaaaccaaac -----
cDNA_rat genomic_hu	5851 1	gggtgtctcagaaaacggccaagggaagcagaaaggctctgggtaacacctg -----
cDNA_rat genomic_hu	5901 1	atggaacattgatcatotataatctgagtcctttatgatcgtgggttttac -----
cDNA_rat genomic_hu	5951 1	aagtgtgtggccagcaacccatctggccaggattcactgttggttaagat -----
cDNA_rat genomic_hu	6001 1	acaagtcatcacagctccccctgtcattatagagcaaaagaggcaagcca -----
cDNA_rat genomic_hu	6051 1	tcgttggggttttaggtggaagtttgaaaotgccctgcactgcaaaagga -----
cDNA_rat genomic_hu	6101 1	actccccagcctagtgttcactgggtcctttatgatgggactgaactaaa -----
cDNA_rat genomic_hu	6151 1	accattgcagttgactcattccagatttttcttgatccaaatggaactc -----
cDNA_rat genomic_hu	6201 1	tgtatataagaagcatcgctccttcagtgaggggcacttatgagtgcat -----
cDNA_rat genomic_hu	6251 1	gccaccagctcctcaggctcagagagaagggtagtgattcttactgtgga -----

ctg	rat	7351	CttTGGAAGCAGAGTCACGGTCCATCCAAATGGAACCTTGGAgatgagga
---	le_hu	890	CtaTGGAAGCAGAATCACAGTCCATAAAATGGAACCTTGGAsattagga
gat	rat	7401	acatccGGCTTTCTGACTCTGCCGACTTCACCTGTGTGGttcggagcgag
---	le_hu	940	atgtgaGGCTTTCAGATTACGCCGACTTTATCTGTGTGGcccgaaatgaa
cca	rat	7451	ggaGGAGAGAGTGTGTGGTAGTGCAGTTAGAAGTCTAGAAATGCTGAG
---	le_hu	990	ggtGGAGAGAGCGTGTGGTAGTACAGTTAGAAGTACTGGAATGCTGAG
IGg	rat	7501	AAGACCAACATTTCAGAAACCCATTCAACGAAAaagtcacgcgccaagctg
IGt	le_hu	1040	AAGACCGACATTTAGAAATCCATTTAATGAAAAaatagttgccagctgg
3CA	rat	7551	gcaagcccgtagCACTGAAGTGTCTGTGGATGGGAACCCCCACCTGAA
3CA	le_hu	1090	gaaagtccacAGCATTGAATTGCTCTGTGATGGTAACCCACCACCTGAA
ATT	rat	7601	attacCTGGATCTTACCTGACGGCACACAGTTTGCTAACAGACCAacaaa
ACT	le_hu	1140	ataatCTGGATTTTACCAAATGGCACACGATTTTCCAATGGACCAcaaaag
CC	rat	7651	ttccccGTATCTGATGGCAGGCAATGGCTCTCTCATCCTTTACAAAgcaa
CC	le_hu	1190	ttatcagTATCTGATAGCAAGCAATGGTTCTTTTATCATTTCTAAAacaa
CT	rat	7701	ctcggaacaagtCAGGGAAGTATCGCTGTGCAGCCAGGAATAAGGTGGC
CA	le_hu	1240	ctcgggaggatgCAGGAAAATATCGCTGTGCAGCTAGGAATAAGGTGGC
CA	rat	7751	TAcacgagaaactcatcctgttagagATTGGGCAGAAGCCAGTCATTCT
CA	le_hu	1290	TAtattgagaaattagtoaatattagaaATTGGCCAGAAGCCAGTTATTCT
TG	rat	7801	GACATACGAACACAGGGatggtgaagagcgTCAGTGGGAACCGTTATCAC
TG	le_hu	1340	TACCTATGCACCAGGGAacagtaaaaggoatCAGTGGAGAATCTCTATCAC
TC	rat	7851	TGCATTGTGTGTCTGATGGgATCCCCAAGCCAAATGTCAAGTGGACTACA
CT	le_hu	1390	TGCATTGTGTGTCTGATGGaATCCCTAAGCCAAATATCAAATGGACTATG
AC	rat	7901	CCGGGTGGccatGTAATCGACAGGCCTCAAGTGGATGGAAAATACATAcT
AC	le_hu	1440	CCAAGTGGttatGTAGTAGACAGGCCTCAAATTAATGGGAAATACATatT
GA	rat	7951	GCATGAAAATGGCAGCTGGTCATCAAAGCAACAACAGCtcacgaccaAG
GA	le_hu	1490	GCATGACAATGGCACCTTAGTCATTAAAGAGCAACAGCttatgacagAG
aa	rat	8001	GAAATTATATCTGTAGGGCTCAAAACAGTGTGGCCAggcagttattagc
at	le_hu	1540	GAAACTATATCTGTAAGGCTCAAAATAGTGTGGTCatacaotgattact
TT	rat	8051	gtgtCAGTGATGGTTGTGGCCTACCCCTCCCGAATCATAAActacotACC
TT	le_hu	1590	gttcCAGTAATGATTGTAGCCTACCCCTCCCGAATTACAAAtogtccACC
TA	rat	8101	CAGGAACATGCTCAGGAGGACAGGGGAAGCCatgCAGCTCCACTGTGTGG
TA	le_hu	1640	CAGGAGTATTGTCAACAGGACAGGGGCAGCctttCAGCTCCACTGTGTGG
TC	rat	8151	CCTTGGGAATCCCCAAGCCaaAAGTCACCTGGGAGACGCCAAGACACTCC
C	le_hu	1690	CCTTGGGAGTTCCCAAGCCagAAATCACATGGGAGATGCCTGACCACTCC
C	rat	8201	CTGCTCTCAAaagcaacagcaagaaacccCATAGAAGTGAGATGCTTCA
C	le_hu	1740	CTTCTCTCAAcggcaagttaaagagaggacaCATGGAAGTGAGCAGCTTCA
TT	rat	8251	CCCACAAGGTACGCTgTTCATTTCAGAATCTCCAAACCTCGGATTCCGGag
TT	le_hu	1790	CTTACAAGGTACCCTaGTCATTTCAGAATCCCCAAACCTCCGATTCTGGga
TT	rat	8301	tcTATAAGTGACAGAGCTCAGAACCTACTTGGgactGATTACGCAACAACT
TT	le_hu	1840	taTACAAATGCACAGCAAAGAACCCACTTGGtagTGATTATGCAGCAACG
TT	rat	8351	TACATCCAGGTACTCTGACAGGAagggggagactaaaattoaacagaaagt
TT	le_hu	1890	TATATTCAGTAATCTGACATGAAataataaagtcacaacatctgggca

cDNA_rat	8401	ccacatccacagggTTTATTTTGGGAAGAAGTTTAATCAAAGGCAGCCA
genomic_hu	1940	gaa-----TTTATTTTGGGAAGAAGTTTAATCAAAGGCAGCCA
cDNA_rat	8451	TAGGCATGTAAATGAgctcGAATACATTACAGTATTAAATTACAATGG
genomic_hu	1979	TAGGCATGTAAATGAatttGAATACATTACAGTATTAAATTACAATGA
cDNA_rat	8501	ACATGCgatga-----GACTTGTAATGAAAGCATTGTGAAGTGAaaccg
genomic_hu	2029	ACATGCaaaataaaaagGACTTGTAATAAATGCATTATGAAGTGAatgata
cDNA_rat	8546	agtctctg-----TGGATCTCAAAGCAAACCTCTTAAGTAAAGGCACCTTg
genomic_hu	2079	otgatttatttaattGGATCTCAAACAAACTTTTAACTTAAGGCACCTTt
cDNA_rat	8591	ATTTTGCCAACAAATAATAACAAACattaagagaaaaaaatgatcCACTA
genomic_hu	2129	ATTTTGCCAACAAATAACAATAAACaaacattgaaacggtt----CACTA
cDNA_rat	8641	CGAAATAACAAACGGCTAATGCACCTGAATTtctcagtaaaaagacotttc
genomic_hu	2175	TAAATAACAAATGGCTAATGTACCTGAATTtttcagtaaaaaaatgaac
cDNA_rat	8691	tctcgctaacagttgCCAGCTGCCTCGTGTCTGTTTCCTACCAATGTCAC
genomic_hu	2225	ttctaata-----CCAGTTGCCTAGTGTCCACCTCCTATCAATGTTAC
cDNA_rat	8741	AAACatogcacacaggggGAATGGAGTCAACGGGAAAGATTAAGTTTGCG
genomic_hu	2268	AAGCATggcactca----GAACAGAGACAATGGAAATATTAATCTGCR
cDNA_rat	8791	GTCTgtgtaaatctca-----ATGTACAAATATTCTGtcncTG
genomic_hu	2314	ATCtttatgatgtaaatttaccatcctgATGTATAAATATTTTG----TG

exon1 (2342-2397)

cDNA_rat	8791	GTCTgtgtaaatctca-----ATGTACAAATATTCTGtcncTG
genomic_hu	2314	ATCtttatgatgtaaatttaccatcctgATGTATAAATATTTTG----TG
cDNA_rat	8829	GTTTATAAACATTTTGATAAAACCGAAAAAAAAAAAAAAAAAAAAA
genomic_hu	2360	GTTTATAAATTTTTTGCTAAACCTACAGAAATAAGcactgaactgtc

(Genomic human OCP: SEQ ID NO:4)
(cDNA rat: SEQ ID NO:5)

Figure 11

Exon/Intron No.	Exon start	Exon end	Exon length	Intron length	Remarks
1	1	208	208	69	No valid splice site found upstream this exon
2	277	429	153	18	
3	447	485	39	1561	
4	2046	2244	199	1351	
5	3595	3724	130	3254	
6	6978	7359	382	4123	
7	11482	14903	3422	38	
8	14941	15307	367	51	
9	15358	15825	468	1039	
10	16864	~17760	~897	?	Last exon might be complete
11	?(1)	2317	~231 7	25	Exon is not complete and start site is not known
12	2342	2397	56	-	

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

Fig 12.

608 Human translated nucleotide sequence (ORF)

ATGAAGGTAAAAGGCAGAGGAATCACCTGCTTGGTCTCCTTTGCTGT
GATCTGCCTGGTCGCCACCCCTGGGGGCAAGGCCTGCTCGCCGCTGTG
CCTGTTATATGCCTACGGAGGTACACTGCACATTTTCGGTACCTGACTTCCA
TCCCAGACAGCATCCCGCCCAATGTGGAACGCATCAATTTAGGATACAAC
AGCTTGGTTAGATTGATGGAACAGATTTTCTGGCCTGACCAAAGTGA
GTTACTCATGCTTCACAGCAATGGCATTACACAATCCCTGACAAGACCTT
CTCAGATTTGCAGGCCTTGCAGGTCTTAAAAATGAGCTATAATAAAGTCC
GAAAACTTCAGAAAAGATACTTTTATGGCCTCAGGAGCTTGACACGATTG
CACATGGACCACAACAATATTGAGTTTATAAACCCAGAGGTTTTTTATGG
GCTCAACTTTCTCCGCTGGTGCCTTGAAGGAAATCAGCTCACTAAGCT
CCACCCAGATACATTTGTCTCTTTGAGCTACCTCCAGATATTTAAATCTC
TTTCATTAAGTTCTTACTTGTCTGATAACTTCTGACCTCCCTCCCTCAA
GAGATGGTCTCCTATATGCCTGACCTAGACAGCCTTTACCTGCATGGAAA
CCCATGGACCTGTGATTGCCATTTAAAGTGGTTGTCTGACTGGATACAGCC
AGATGTAATAAAATGCAAAAAGATAGAAGTCCCTCTAGTGCTCAGCAGT
GTCCACTTTGCATGAACCTTAGGACTTCTAAAGGCAAGCCGTTAGCTATG
GTCTCAGCTGCAGCTTTCCAGTGTGCCAAGCCAACCATGACTCATCCCTG
AAATCAAAGAGCCTGACTATTCTGGAAGACAGTAGTTCTGCTTTTCATCTCT
CCCCAAGGTTTCATGGCACCCCTTTGGCTCCCTCACTTTGAATATGACAGAT
CAGTCTGGAATGAAGCTAACATGGTCTGCAGTATTCAAAAGCCCTCAAG
GACATCACCCATTGCATTCACTGAAGAAAATGACTACATCGTGCTAAATA
CTTCATTTCAACATTTTGGTGTGCAACATAGATTACGGTCACATTCAGC
CAGTGTGGCAAATTTTGGCTTTGTACAGTGATTCTCCTCTGATACTAGAAA
GGAGCCACTTGCTTAGTGAAACACCGCAGCTCTATTACAAATATAAACAG
GTGGCTCCTAAGCCTGAAGACATTTTACCAACATAGAGGCAGATCTCAG
AGCAGATCCCTCTTGGTTAATGCAAGACCAAATTCCTTGCAGCTGAACA
GAACTGCCACCACATTCAGTACATTACAGATCCAGTACTCCAGTGATGCT
CAAACTCACTTTACCAAGAGCAGAGATGAGGCCAGTGAAACACAAATGGA
CTATGATTTCAAGGGATAACAATACTAAGCTGGAACATACTGTCTTGGTA
GGTGGAAACCGTTGGCCTGAACTGCCAGGCCAAGGAGACCCACCCCA
CGTGGATTGGCTTCTAGCTGATGGAAGTAAAGTGAGAGCCCTTATGTCA
GTGAGGATGGACGGATCCTAATAGACAAAAGTGGAATAATGGAAGTCCA
GATGGCTGATAGTTTTCACACAGGCGTATATCACTGTATAAGCAGCAATT
ATGATGATGCAGATATTTCTACCTATAGGATAACTGTGGTAGAACCTTTGG
TCGAAGCCTATCAGGAAAATGGGATTCATCACACAGTTTTCATTGGTGAA
ACACTTGATCTTCCATGCCATTCTACTGGTATCCAGATGCCTCTATTAGC
TGGGTTATTCCAGGAAACAATGTGCTCTATCAGTCATCAAGAGACAAGAA
AGTTCTAAACAATGGCACATTAAGAATATTACAGGTCAACCCGAAAGACC
AAGGTTATTATCGCTGTGTGGCAGCCAACCCATCAGGGGTTGATTTTTGA
TTTTCCAAGTTTCAGTCAAGATGAAAGGACAAAGGCCCTTGGAGCATGAT
GGAGAAACAGAGGGATCTGGACTTGATGAGTCCAATCCTATTGCTCATCT
TAAGGAGCCACCAGGTGCACAACTCCGTACATCTGCTCTGATGGAGGCTG
AGGTTGGAAAACACACCTCAAGCACAAAGTAAGAGGCACAACTATCGGGA
ATTAACACTCCAGCGACGTGGAGATTCAACACATCGACGTTTTAGGGAGA
ATAGGAGGCATTTCCTCCCTCTGCTAGGAGAATTGACCCACAACATTGG
GCGGCACTGTTGGAGAAAAGCTAAAAAGAAATGCTATGCCAGACAAGCGAG
AAAAATACCACAGTGAGCCCAACCCCAAGTGGTCAACCAACTCCCAACATA
CCTGGTGAAGAAGACGATTCTCAGGCATGCTCGCTCTACATGAGGAATT

TATGGTCCCGGCCACTAAAGCTTTGAACCTTCCAGCAAGGACAGTGA CTGACTCCAGAACAATATCTGATAGTCCTATGACAAACATAAATTATGGC ACAGAATTTCTCTCCTGTTGTGAATTCACAAATACTACCACCTGAAGAACCC ACAGATTTCAAACGTGTCTACTGCTATTAATAACTACAGCCATGTCAAAGAA TATAAACCCCAACCATGTGCAAGCCAAATACAAGGCACAACCAATCAACATT CATCCACTGTCTTTCCACTGTCTACTTTGGAGCAACTGAATTTTCAGGACTCTG ACCAGATGGGAAGAGGAAGAGAGCATTTCCAAAGTAGACCCCAATAAC AGTAAGGACTATGATCAAAGATGTCAATGTCAAAATGCTTAGTAGCACC CCAACAACATAATTATTAGAGTCAGTAATAACCAAAATAGTCATCAGACA TCTGTAAGAGAAGATGAGTGAACCCAGGCACAATCACTTCTATTCTCACAC TACTCAAATACTTAGCACCTCCACGTTCCCTTCAGATCCACACACAGCTGC TCATTCTCAGTTTCCGATCCCTAGAAATAGTACAGTTAACATCCCCTGTT CAGAGCCTTTGGGAGGCAGAGGAAATTTGGCGGAAGGGGGCGGATTATC AGCCCCATAAGAACCTCCAGTTTCTGCGACGGCATAGATACAGCAATTTTCAG GTCAACAACCCAGAGGTTCTTCTGAAAAAAGCACTACTGCAATTTCCAGCCA CAGTGCTCAATGTGACATGTCTGTCTCTGCTTCCCAGGGAGAGGGCTCACC A CTGCCACAGCAGCATTGTCTTTTCCAAGTGCTGCTCCCATCACCTTCCCCA AAGCTGACATTTGCTAGAGTCCCATCAGAAGAGTCTACAACCTAGTCCAG AATCCCACTATTACTACTGTGAGAACAACCCAGGTGAGAGAAAAACAACCC CACAATAAAAAATTTTCAGGACTGAAATTTTCCCAAGTGAAGTCCCAACTGGTG CAGTCATGACATATGCTCCAACATCCATACCCATGGAAAAAACTCACAAA GTAAACGCCAGTTACCCACGTGTGTCTAGCACC AATGAAGCTAAAAGAGA TTCAGTGATTACATCGTCACTTTTCAGGTGCTATCACC AAGCCCAATGAC TATTATAGCCATTACAAGGTTTTCAAGAAGGAAAAATCCCTGGCAACAGA CACTTTGTAAATAACCATACAACCAAAAAAGGCAGATTAAAGGAATCAACATAAA GTTAGTTTACAAAAAAGCACAGCTGTGATGCTTCTTAAAAACATCTCCTGCT TTACCACAGAGACAAAGTTCCCTTTTCCATTTACCCACACTTTCAACAAGT GTGATGCAAAATTCATCTAATACCTTTGACTACCGCTCACCACTACCGACC AAAACACACAATCTGGGAAGCTTTCCACAAAGAAGAGGCTTCCCTTCCC ACCCTTAACCCATAGTCTTCTGATTATAAGCAAAAGACTCAAGTACAA AAAGCATCATATCAACGCAAAACAGCAATACCAGCAACAACCTCCTACCTTC CTTGCATCTGTATCACTTATGAAACCCAAACAGAGAGATCTAGAGCACA AACAATACAAAGAGAGAACAGGAGCCTCAAAAAGAAACAGGACTGACCCA AACATCTCTCCAGACCAAGAGTTCTGGCTTCTACTACACCCCACTGCTATGACA CCTCTGCTCTGGCATTCACTTATCCCCACCAGAAAAACAACCTGGGATT TCAAGCACAAATCAGTTTTCATTCAAGAACTCTTAATCTGACAGATGTGATT GAAGAACTAGCCCAAGCAAGTACTCAGACTTTGAAGAGCACAATTGTCTC TGAACAACACTTTGTCCAGCAAAATCACACCAAGTAGTACCACAACCTAGGAAA G CATCATTAGACACTCCCCATACCACCTATTCTTGAGCAGCAGTGCTACTTAA TGCCAGTTCCCATCTCCCCCTCCTTTACTCAGAGAGCAGTTACTGACACAC GTGGCGACTCCCATTTCGGCTTATGACAAATACAGTGGTCAAGCTGCAC GAATCCTCAAGGCACAATCTCCAAATGCCAAGTTCACAATTGGAACCACT CACTTCATCTACCTCTAATCTGTTACATTCTACTCCCCAGCCAGCAATAAC AACAGTTAAATCACAGAAATCCAAATTAACCTCATCTCCTGGGCAGAAAT ACCAATTTTGGCACAAACCATACATCAGACATTGCTGAAAAGGCCAAAAAG CCAGAAGTAAGCATGTTGGCTACTACAGGCCTGTCCGAGGCCACCACTCT TGTTTCAGATTGGGATGGACAGAAGAACACAAAGAAGAGTGACTTTGATA AGAAACCGATTTCAAGAAGCAACAACCTTCCAACTCTTCCCTTTGACTCTT TGCTTAGGTATATATTTGAAAAAGCCAGGATAGTTGGAGGAAAAAGCTGCA AGTTTACTATTCCAGCTAACTCAGATGCCTTCTCTCCCTGTGAAGCTGTT

GGGAATCCCCTGCCACCACCATTCATTGGACGACAGAGTTTCAGGACTTGATTGTA
TCTAGAGGAAACCAGAATAGCAGGGTCCAGGTTCTCCCCAATGGTACCCT
GTCCATCCAGAGGGTGGAAATTCAGGACCGCGGACAGTACTTGTGTTCCG
CATCCAATCTGTTTGGCACAGACCACCTTCATGTCACCTTGTCTGTGGTTT
CCTATCTCTCCAGGATCCTGGAGAGACGTACCAAAAGAGATCAGATTCTAT
TCCGGAAGCACTGTGGAATCTGAAGTGCAGAGCAGAAGGTAGGCCAAGCC
CTACAGTTACCTGGATTCTTGCAAACCAAACAGTTGTCTCAGAATCATCCC
AGGGAAGTAGGCAGGCTGTGGTGACGGTTGACGGAACATTGGTCTCTCCAC
AATCTCAGTATTTATGACCGCTGGCTTTTACAAATGTGTGGCCAGCAACCCA
GGTGGCCAGGATTCCTGCTGGTTTAAAAACAAGTCATTGCAGCACCAACC
TGTTATTCTAGACCAAAGGAGGCAAGTCATTGTAGGCACCTTGGGGTGAAA
GTTTAAAACTGCCCTGTACTGCAAAAGGAACTCCTCAGCCCAGCGTTTACT
GGGTCTCTCTGATGGCACTGAAGTGAAACCATTACAGTTTACCAATTCCA
AGTTGTTCTTATTTTCAAATGGGACCTTTGTATATAAGAAACCTAGCCTCTT
CAGACAGGGGCACTTATGAATGCATTGTCTACAGTTCCACTGGTTCCGGAG
CGAAGAGTAGTAATGCTTACAATGGAAGAGCGAGTGACCAGCCCCAGGA
TAGAAGCTGCATCCCAGAAAAAGACTGAAGTGAATTTGGGGACAAATTA
CTACTGAACTGCTCAGCCACTGGGGAGCCCCAAACCCCAATAATGTGGAG
GTTACCATCCAAGGCTGTGGTGCAGCAGTGGAGCTGGATCCACGCTTACCC
CTAATGGATCCCTGTTTATTGGATCAGTAACAGAAAAAGACAGTGGTGTC
TACTTGTGTGTGGCAAGAAACAAAATGGGGGATGATCTGATACTGATGCA
TGTTAGCCTAAGACTGAAACCTGCCAAAATTGACCACAAGCAGTATTTTA
GAAAGCAAGTGCTCCATGGGAAAGATTTCCAAGTAGATTGCAAAGCTTCC
GGCTCCCCAGTGCCAGAGATATCTTGGAGTTTGCCTGATGGAACCATGAT
CAACAATGCAATGCAAGCCGATGACAGTGCCACAGGACTAGGAGATAT
ACCCTTTTCAACAATGGAACCTTTATACTTCAACAAAGTTGGGGTAGCGGA
GGAAGGAGATTATACTTGCTATGCCCAGAACACCTAGGGAAAGATGAA
ATGAAGGTCCACTTAACAGTTATAACAGCTGCTCCCCGGATAAGGCAGAG
TAACAAAACCAACAGAAGATAACAAAGCTGGAGACAGACGCTGCTCTTGACT
GTGAGGTCACTGGGGATCCCAACAAAAAATATTTGGTTGCTGCCTTCC
AATGACATGATTTCTTCTCCATTGATAGGTACACATTCATGCCAATGGG
TCTTTGACCATCAACAAAGTGAAACTGCTCGATTTCTGGAGAGTACGTATG
TGTAGCCCGAAAATCCCAGTGGGGATGACACCAAAATGTACAAAATGGATG
TGGTCTCTAAACCTCATTAACTCAATGTCTGTATACAAACAGAACTGTTA
TTAAAGCCACAGCTGTGAGACATTCCAAAAAACACTTTGACTGCAGAGCT
GAAGGGACACCATCTCCTGAAGTCATGTGGATCATGCCAGACAATATTTT
CCTCACAGCCCCATACTATGGAAGCAGAATCACAGTCCATAAAAAATGGAA
CCTTGGAATATTAGAAATGTGAGGCTTTTCAGATTTCAGCCGACTTATCTGTG
TGGCCCGAAATGAAAGGTGGAGAGCGTGTTGGTAGTACAGTTAGAAGTA
CTGAAATGCTGAGAAGACCGACATTTAGAAATCCATTTAATGAAAAAT
AGTTGCCAGCTGGGAAAGTCCACAGCATTGAATTGCTCTGTTGATGGTA
ACCCACCACCTGAAATAATCTGGATTTTACCAAATGGCACAGCATTTTCCA
ATGGACCACAAAGTTATCAGTATCTGATAGCAAGCAATGGTTCTTTATCA
TTTCTAAAAACAACCTCGGGAGGATGCAAGGAAAAATTCGCTGTGCAGCTAGG
AATAAAGTTGGCTATATTGAGAAATTAGTCATATTAGAAATTGGCCAGAA
GCCAGTTATTCTTACCTATGCACCAGGGACAGTAAAAGGCATCAGTGGAG
AATCTCTATCACTGCATTGTGTGTCTGATGGAATCCCTAAGCCAAATATCA
AATGGACTATGCCAAGTGGTTATGTAGTAGACAGCGCTCAAATTAATGGG
AAATACATATTGCATGACAATGGCACCTTAGTCATTAAGAAAGCAACAGC
TTATGACAGAGGAAACTATATCTGTAAGGCTCAAAATAGTGTGGTCATA

Figure 13

A

Region			Region Length	% identity	% positives	% gaps
General	Rat	Human				
1-655	1-655	1-653	655	76	86	0
656-726	656-726	654-724	71	46	62	0
727-779	727-779	725-777	53	77	86	0
780-1634	780-1617	778-1612	655	38	53	4
1635-end	1618-end	1613-end	980	74	85	0
	Total	Alignment	2614	62	74	1

B

Region			% identity	% positives	% gaps
Rat	Mouse	Length			
1-238	1-238	238	91	92	1

C

Region			Region Length	% identity
General	Rat	Human		
1-1965	1-1965	1-1965	1965	83
1966-2178	1966-2178	1966-2178	213	86
2179-2337	2179-2337	2179-2337	159	86
2338-4893	2338-4893	2338-4863	2565	63
4894-7833	4852-7791	4864-7761	2940	80
	Total	Alignment	7842	80

D

Region			Region Length	% identity
General	Rat	Mouse		
1-720	1-718	1-720	720	93

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 14

rat_cDNA human_5+3_corrected mus_cDNA_5	CGAGAGACGACAGAGGTTACGGCTGCGAGAAGACGACAGAGGGTCCAGAAAAGGAAA -----CAGAGGGTCCAGGAAA-GGAAA
rat_cDNA human_5+3_corrected mus_cDNA_5	GTGCTGGAGGGGAGTGGGACAAAAGCAGCGACCAAGTGAATGTCACTTCAGTGACTGAG -----GTACTGGAGGGGAGTTGGGACAAAAGCAGCGACCAAGGGAACATCGCTTCAGTGACTGAA
rat_cDNA human_5+3_corrected mus_cDNA_5	GCCAGGCAAAAAGCGCGGGAAGGATTTTGTGTAGCTTGGGACCCCTTCATAGACACTGAT -----GCCAGGCAAAAAGGAGCGGGAAGGATTATATGTAGCCTGGGACGCTTCATAAACACTGAT
rat_cDNA human_5+3_corrected mus_cDNA_5	GACACGTTTACGCAAAATA-GAAATTTGAGGAGAAACGCCTGGGCCTTCGGAAAG---GA -----GACGTGTTTGTGCAAGCAAGCAATTTGAGGAGAAACGCCTGGGACGTCGGAAAGAAAGGA
rat_cDNA human_5+3_corrected mus_cDNA_5	GTGATTGATTAGTACTTTCAGTGTAGGTGACTTTAAGGAGAACTAACTAATGTATACTA -----GTGATCGATTAGTACTTGTAGTGTAGGTGAGTTT---GAGAACTAACTAACTATACTA
rat_cDNA human_5+3_corrected mus_cDNA_5	TTGAGGGAGGAGGAAGAGCATTACAGAGTTTCCAGCAGCAGCAGGAAAGCTTTGGTTAAT -----TTGAGGGAGGAAGGAAGAGCATT-----CCAGCAGCAGCAGGAAAGCTTTGGTTAGT
rat_cDNA human_5+3_corrected mus_cDNA_5	TTGGAAATGGATGATAGCATTAAATAACAGAAGCGCCTCCAGGTCTCTGAAGCTTCAGT -----TTGGAAATGTATGATACCATTAAATAACAGAAGCGCCTCCAGTTCTCTGAAGAGTCAGT
rat_cDNA human_5+3_corrected mus_cDNA_5	CCCCAGCTGAAAGCCAGAAAGACTAAGCCCTAAGCCCTTTTGATCCCTTTGGAAGCA -----CCCCAGCT-----A-GTGTAAGCCTACTAAGCCCTTTTGCTCCCGTTGGAAGCA
rat_cDNA human_5+3_corrected mus_cDNA_5	AAGAAGCTTTCCTTCCCTGGGTGAAGACTCTCCTCAGAAGATTTCCTGTCTCTGCCTATG -----AAGAAGCTTTCCTTCAATCAGGTGAAGGCTCTCCTCAGAAGATTTCCTGTCTCTGCCTATG
rat_cDNA human_5+3_corrected mus_cDNA_5	TTACAAGAGGAATCAAAOCCAAGACAGAAGAGCTCAGGATGCAGGTGAGAGGCAGGGAG -----ATGAAGGTAAAGGCAGAGGAA TTACAAGAGGATTCAAAAGCAAGACAGAAGAGCTCAGGATGCAGAAGAGAGGCAGGGAG *** ** * ***** *
rat_cDNA human_5+3_corrected mus_cDNA_5	TCAGCGGCTTGTGTATCTCCCTCACTGCTGTCTGCCTGGTGGTCAOCCCTGGGAGCAGGG TCACTGCTGTGCTGCTCTCCTTTGCTGTGATCTGCCTGGTGGTCAOCCCTGGGAGCAGGG TCAGCTGCTGTGCTGATCTCCTCACTGCCATCTGCCTGGTGGTCAOCCCTGGGAGCAGGG *** * ***** * ***** * ***** * ***** * ***** *
rat_cDNA human_5+3_corrected mus_cDNA_5	OCTGTCTCGCGCTGTGCTGTATGTGCCACAGAGGTGCAGTGTACATTTGGGTACC CCTGTCTCGCGCTGTGCTGTATGTGCCACAGAGGTGCAGTGTACATTTGGGTACC TCTGTCTCGCGCTGTGCTGTATGTGCCACAGAGGTGCAGTGTACATTTGGGTACC ***** * ***** * ***** * ***** * ***** * ***** *
rat_cDNA human_5+3_corrected mus_cDNA_5	TGACCTCCATCCCGATGG-CATCCCGGCCAATGTGGAACGAATAAATTTAGGATATAAC TGACTTCCATCCCGACAG-CATCCCGGCCAATGTGGAACGAATAAATTTAGGATATAAC TGACCTCCATCCCGACAGGGCATCCCGGCCAATGTGGAACGAATAAATTTAGGATATAAC *** * ***** * ***** * ***** * ***** * ***** *

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AGCCTTACTAGATTGACAGAAAAAGACTTTGATGGCCTGAGCAAACCTGGAGTTACTCATG
AGCTTGGTTAGATTGATGGAAACAGATTTTCTGGCCTGACCAAACCTGGAGTTACTCATG
AGCCTCACTAGATTGACAGAAATGACTTTTCTGGCCTGAGCAGACTGGAGTTACTCATG
*** * ***** ** * ***** *

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CTGCACAGTAATGGCATTTCACAGAGTCAGTGACAAGACCTTCTCGGGCTTGCACTCCTTG
CTTCACAGCAATGGCATTTCACACAATCCCTGACAAGACCTTCTCAGATTGTCAGGCCTTG
CTGCACAGCAATGGCATTTCACAGAGTCAGTGACAAGACCTTCTCGGGCTTGCACTCCTTG
** ***** * * ***** *

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CAGGTCTTAAAAATGAGCTATAACAAAGTCCAAATCATTGGAGGATACCTTTCTACGGA
CAGGTCTTAAAAATGAGCTATAATAAAGTCCGAAACCTCAGAAAGATCTTTTATGGC
CAGGTCTTAAAAATGAGCTATAACAAAGTCCAAATCATTGGAGGATACCTTTGTATGGA
***** ** * ***** *

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CTCGGGAGCTTGGTCCGCTTGCACTGGATCACAACAACATTGAATTCATCAACCCCTGAG
CTCAGGAGCTTGACACGATTGCACATGGACCAACAATATTGAGTTTATAACCCAGAG
CTCAGGAGCTTGACCCGCTTGCACTGGATCACAACAACATTGAGTTTATCAACCCGAG
*** ***** * * ***** *

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GCTTTTATGACITACCTCGCTCCGCTTGGTACATTAGAAGGAAACCGGCTCACAAG
GTTTTTATGGGCTCACTTTCTCCGCTGGTGCCTTGGAGGAAATCAGCTCACTAAG
GCGTTTACGGACTCACCTTCTCCGCTTGGTACATCTAGAAGGAAACCGGCTCACAAG
* ***** * * ***** *

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CTCCATCCAGACACATTGTCTCATTAAGCTATCTCCAGATATTTAAACCTCTTTCATT
CTCCACCCAGATACATTGTCTCTTTGAGCTACCTCCAGATATTTAAATCTCTTTCATT
CTCCATCCAGACACATTGTCTCTTTGAGCTATCTCCAGATATTTAAACCTCTTTCATT
***** ** * ***** *

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AAGTACCTGTCTGTCTGATAACTTCTT-GACCTCCCTCCCAAAGAAATGGTCTCCTA
AAGTCTCTATACCTGTCTGATAACTTCTT-GACCTCCCTCCCTCAAGAGATGGTCTCCTA
AAGNACCTGTACTGTATGATAACTTCTT-GACCTCCCTCCCAAAGAAATGGTCTCCTC
*** ** * ***** *

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CATGCCAAACCTAGAAAGCCTGTATTGCAATGGAAACCCATGGACCTGTGACTGCCATTT
TATGCTGACCTAGACAGCCTTTACCTGCAATGGAAACCCATGGACCTGTGACTGCCATTT
TATGCCAAACCTAGAAAGCCTTTACTTGCATGGAAACCCATGGACCTGTGACTGCCATTT
***** ** * ***** *

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AAAGTGGTTGTCTGAGTGGATGCAGGGAACCCAGATATAATAAAATGCAAGAAAGACAG
AAAGTGGTTGTCTGACTGGATACAGGNNNNOCAGATGTAATAAAATGCAAAAAGATAG
AAAGTGGTTGTCCGAGTGGATGCAGGGAACCCAGGTA-ACTATCTTGT----TTGTTTG
***** ** * ***** *

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AAGCTCTTCCAGTCTCAGCAATGTCCCT-TTGCATGAACCCAGGATCTCTAAAGGCA
AAGTCCCTCTAGTCTCAGCAGTGTCCCT-TTGCATGAACCCAGGATCTCTAAAGGCA
TTTCTTTTTTATARKAGTATTTTCTCAATTTCAATTAGATGATATCCCAAAGTC-
* * * * *

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GACCCCTTGTCTATGGTACCATCTGGAGCTTTCTATGTACAAAGCCAAACATTGATCCAT
AGCCGTTAGCTATGGTCTCAGCTGCACTTTCCAGTGTGCCAAGCCAAACATTGACTCAT
-CCCATACCTCCGCCCA-----CTTCCCTACCTACCCATTC-CCATTTTTTGGC
** * * * * * * * * * *

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CAGTGAAGTCAAGAGCCTGGTTACTCAGGAGGACAAATGGATCTGCTCCACCTCACCTC
CCCTGAAATCAAGAGCCTGACTATTCTGGAAGACAGTACTTCTGCTTTCATCTCTCCCC
CCTGGCATTCCTC-----
* * * * *

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AAGATTTATAGAAACCTTTGGCTCCTTGTCTTTGAACATGACANANNNTCTGGAATA
AAGGTTTCATGGCACCTTTGGCTCCTCAGTTTGAATATGACAGATCAGTCTGGAATG

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AGGCCGACATGGTCTGTAGTATCAAAAGCCATCAAGGACATCAACCACTGCATTCACTG
AAGCTAACATGGTCTGCACTATTCAAAAGCCCTCAAGGACATCAACCACTGCATTCACTG

AAGAAAATGACTACATCATGCTAAATGCGTCATTTCCACAAATCTTGTGTGCAGTGTAG
AAGAAAATGACTACATCGTGCTAAATACTTCATTTTCAACATTTTTGGTGTGCAACATAG

ATTATAATCACATCCAGCCAGTGTGGCAACTTCTGGCCTTATACAGTGA CTCTCTGA
ATTACGGTCACATTACAGCCAGTGTGGCAAAATTTGGCCTTGTACAGTGATTCTCTCTGA

TACTAGAAAGGAAGCCCGAGCTTACCGAGACTCCTTCACTGTCCTTCTAGATATAAACAGG
TACTAGAAAGGAGCCACTTGCTTAGTGAAACACCGCAGCTCTATTACAATATAAACAGG

TGGCTCTTAGGCTGAAGACATTTTACCAGCATAGAGGCTGATGTCAGAGCAGACCOCTT
TGGCTCCTAAGCTGAAGACATTTTACCAACATAGAGGCAGATCTCAGAGCAGATCCCT

TTTGGTTCCACAAGAAAAAATGTCTTGAGCTGAACAGAACTGOCACCACACTTAGCA
CTTGGTTAATGCAAGACCAAAATTTCCCTTGAGCTGAACAGAACTGOCACCACATTAGTA

CATTACAGATCCAGTFTTCCACTGATGOTCAAATCGCTTTACCAAGGGCGGAGATGAGAG
CATTACAGATCCAGTACTCCAGTGAATGOTCAAATCACTTTACCAAGAGCAGAGATGAGGC

CGGAGAGACTCAAATGGACCATGATCCTGATGATGAACAAATOCCTAACTGGAAAGCACTG
CAGTGAACACAAATGGACTATGATTTCAAGGGATAACAATACTAAGCTGGAAACATACTG

TCTTGGTTAGGTGGAAACCGTTGGCCTGAAGCTGAGGCAAGGAGACCCCAACCCACACG

TGGAATGGCTTCTAGCTGATGGGAGTAAAGTGAGAGCCCTTACGTTAGCGAGGATGGG
TGGATTGGCTTCTAGCTGATGGGAGTAAAGTGAGAGCCCTTATGTCACTGAGGATGGAG

GAATCCTAATAGACAAAATGGGAAGTTGGAACTGCAGATGGCTGACAGCTTTGATGCAG
GGATCCTAATAGACAAAAGTGGAAAATTGGAACTCCAGATGGCTGATAGTTTTGACACAG

GTCTTTACCACTGCATAAGCACCAATGATGCAGATGCGGATGTTCTCACATACAGGATAA
GCGTATATCACTGTATAAGCAGCAATTATGATGATGCAGATATTCTCACCTATAGGATAA

CTGTGGTAGAGCCCTATGGAGAAAGCACACATGACAGTGGAGTCCAGCACACAGTGGTTA
CTGTGGTAGAAOCTTTGGTCGAAGCCTATCGGAAAATGGGATTCATCACACAGTTTCA

CGGGTGAGAOGCTOGACCTTOCATGCCTTTCCACGGGTGTTCCAGATGCTTCTATTAGCT
TTGGTGAAACACTTGATCTTCCATGOCATTTCTACTGGTATCCAGATGCCTCTATTAGCT

GGATTCTTCCAGGGAACACTGTGTTCTCTAGCCATCAAGAGACAGGCAAATTCTTAAC
GGGTTATTCCAGGAAACAATGTGCTCTATCAGTCATCAAGAGACAAGAAAGTTCTAAAC

rat_cDNA human 5+3 corrected mus_cDNA_5	ATGGGACCTTAAGAATATTACAGGTTACGCCAAAAGATCAAGGTCATTACCAATGTGTGG ATGGCACATTAAAGAATATTACAGGTCACCCGAAAGACCAAGGTTATTATCGCTGTGTGG
rat_cDNA human 5+3 corrected mus_cDNA_5	CTGCCAACCCATCAGGGGCCGACTTTTCCAGTTTAAAGTTTCAGTTCAAAGAAAGGCC CAGCCAACCCATCAGGGGTTGATTTTTCAGTTTCCAGTTTCAGTCAAGATGAAAGGAC
rat_cDNA human 5+3 corrected mus_cDNA_5	AAAGGATGGTTGAGCATGACAGGGAGGCAAGTGGATCTGGACTTGGAGAACCCAACTCCA AAAGGCCCTTGGAGCATGATGGAGAAACAGAGGGATCTGGACTTGATGATCCAACTCCTA
rat_cDNA human 5+3 corrected mus_cDNA_5	GTGTTTCCCTTAAGCAGCCAGCATCTTTGAACTCTCTGCATCAGCTTTGACAGGTCAG TTGCTCATCTTAAGGAGCCACAGGTGCACAACTCCGTACATCTGCTCTGATGGAGGCTG
rat_cDNA human 5+3 corrected mus_cDNA_5	AGGCTGGAAACAAGTCTCCGCTGTACATAGGAAGAACAAACATAGAGACTTAATACATC AGGTTGGAAACACACCTCAAGCACAACTAAGAGGCACAACTATCGGGAATTAACACTCC
rat_cDNA human 5+3 corrected mus_cDNA_5	GGCGGCTGGGGATTCCACGCTCCGGGCAITCAGGGAGCATAGGAGGCAGCTCCCTCTCT AGCGACGTGGAGATTCAACACATCGACGTTTATAGGAGAAATAGGAGGCATTTCCCTCCCT
rat_cDNA human 5+3 corrected mus_cDNA_5	CTGCTCGGAGAATTGACCCGCAACGCTGGGCGACACTTCTAGAAAAGCCAAAAGAAAT CTGCTAGGAGAATTGACCCACAACTTGGGCGGCACTGTTGGAGAAAGCTAAAAGAAATG
rat_cDNA human 5+3 corrected mus_cDNA_5	CTGTGCCAAAAAGCAAGAAATACCAAGTAAGCCAGTGCCACTGGCTGTTCCTCTG CTATGCCAGACAGCGAGAAATACCAAGTGAGCCGACCCCACTGGTCAACCCAACTCC
rat_cDNA human 5+3 corrected mus_cDNA_5	TGGAACTCACTGACGAGGAAAGGATGCTCTGGCATGATTCCTCCAGATGAAGAATTCA CAACATACCTGCTGAAGAAGACGATTCCTCAGGCATGCTCGCTCTACATGAGGAATTGA
rat_cDNA human 5+3 corrected mus_cDNA_5	TGTTCTGAAAACCTAAGGCTTCTGCTGCTCCAGGAAGGTCAACCACTGCTGACTCTGGAC TGTTCCCGCCACTAAAGCTTTGAACCTTCCAGCAAGGACAGTGACTGCTGACTCCAGAA
rat_cDNA human 5+3 corrected mus_cDNA_5	CAGTAAATCATGGTTTATGACGAGTATAGCTTCTGGCACAGAAGTCTCAACTGTGAATC CAATATCTGATAGTCCTATGACAAACATAAATTATGGCACAGAAGTCTCCGTTGTGAATT
rat_cDNA human 5+3 corrected mus_cDNA_5	CACAAACACTACAATCTGAGCACCTTCTGATTCAAATTATTTAGTGTAAACAAAGGTA CACAAATACTACCACTGAAGAACCACAGATTCAAAGTGTCTACTGCTAATTAACAACTA
rat_cDNA human 5+3 corrected mus_cDNA_5	CAGCTGTGACAAAGAGTATGAACCCATCCATAGCAAGCAAATAGAAGATACAACCAACC CAGCCATGTCAAAGAATATAACCCCAACCATGTCAAGCCAAATACAGGCACCAACCAATC
rat_cDNA human 5+3 corrected mus_cDNA_5	AAAACCCAATCATTTATCTTTCCATC-----AGTAGCTGAAATTCGAGATTCTG-CT AACATTCATCCACTGCTTTCCACTGCTACTTGGAGCACTGAATTTCAGGACTCTGACA

GTCACCTGAGCATGCCTGGGACCATCCAACTGGGAAAGATTCA GTGGAAACACACCA
GTTACCCACGTGTGTCTAGCACCAATGAAGCTAAAAGAGATTCA GTGATTACATGTCAT

rat_cDNA human_5+3_corrected mus_cDNA_5	TTCCGAGCCCCCTCAGCACACCCCTCAATAACAA-----CAAGCACAAAATTTCTCAAAGA TTTCAGGTGCTATCAOCAAAGCCCAATGACTATTATAGOCATTACAGGTTTTCAAGAA -----
rat_cDNA human_5+3_corrected mus_cDNA_5	GGAAAACTCCCTTGCAACAGATCTTTGTAAATAACCAAGAAGGAGGGGATGTTAAAGA GGAAATTCCTGGCAACAGAACTTTGTAAATAACCAATAACCAAAAGGCAGATTAAAGA -----
rat_cDNA human_5+3_corrected mus_cDNA_5	ATCCATATCAATTGGTTTACAAAAGAACCCAGCCGCAAGGTTCCCAAAATAGCTCCTC ATCAACATAAAGTTAGTTTACAAAAGCAACAGCTGTGATGCTTCCTAAAACATCTCCTG -----
rat_cDNA human_5+3_corrected mus_cDNA_5	TTTATCCACAG-GTCAGAGTTCCCCCTCAGATTCTACAACCTCTCTTGACAAGTCCGCA CTTTACC-ACAGAGACAAAGTTCCCTTTCCATTTCAACCACTTTCAACAAGTGTGATG -----
rat_cDNA human_5+3_corrected mus_cDNA_5	CCAGCTCTGTCTACAACAATGGCTGCCACTCAGAACAAGGGCACTGAAGTAGTATCAGGT CAAATTCATCTAATACCTTGACTACCGCTCACCACACTACGAACAAA--ACACACAA-T -----
rat_cDNA human_5+3_corrected mus_cDNA_5	GCCAGAAGTCTCTCAGCAGGGAAGAAGCAGCCCTTCAOAACTCCTCTCCAG-TGCTTCC CCTGGAAGTCTTCCAACAAAGAAGGAGCTTCCCTTC-CCACCCCTTAACCTATGCTTCC -----
rat_cDNA human_5+3_corrected mus_cDNA_5	TAGCOCATAAGCAAGAGATCTAATACATTAACTTCTTGTCAACGGAAACCCCCAGT TAGTATTATAAGCAAGACTCAAGTACAAAAGCATCATATCAACGCAACAGCAACGGC -----
rat_cDNA human_5+3_corrected mus_cDNA_5	GACAAGTCTACTGCTACTGCTATCTGTCTATTATGTCTGAAACCAACGAACAGATCCAA AACAACTCTACCTTCCCTGCTATCTGTCTACTTATGAAACCAACAGAGAGATCTAG -----
rat_cDNA human_5+3_corrected mus_cDNA_5	AGAAGCAAAGACCAATATAAGGGGCTCG---GAAGAACAAGAACGCAACACCCAC AGCACAACAAATACAAAGAGAAGGAOCTCAAAAGAGAAGCAGGACTGACCCAAACATCTC -----
rat_cDNA human_5+3_corrected mus_cDNA_5	CCCCAGGCAGGTTTCTGGCTATAGTGCACTACTCAGCTCTAACACAGCTGATAACCCCTT TCCAGACCAGAGTTCTGGCTTCACTACACCACTGCTATGACNACCTCCTTNGCTCTNNN -----
rat_cDNA human_5+3_corrected mus_cDNA_5	GGCTTTCAGTCAATCCCCACGACAAGATGATGGTGGAAATGTAAGTGCAAGTTGCTTATCA NGCATTCATCTCATTCOCCACCAGAAAACACAACCTGGGATTTCAAGCACAATCAGTTTTCA -----
rat_cDNA human_5+3_corrected mus_cDNA_5	CTCAACAACCTCTCTCTGGCCA---TAACTGAAGTGTGAG-AAGTAC--ACCCAGAC TTCAAGAACTCTTAATCTGACAGATGTGATTGAGAAGTACCCAGCAGTACTCAGAC -----
rat_cDNA human_5+3_corrected mus_cDNA_5	TTTGGGAATACACAGCTTTGGAACAACGTTGTTGAGCAATCACAGGAGAGTACCAC TTTGAAGAGCACAATTGCTTCTGAAACAACCTTTGTCCAGCAATCACACCAGAGTACCAC -----
rat_cDNA human_5+3_corrected mus_cDNA_5	AGTGAAGAGAGCTCA---GACAC---ACCAACCACTCCTCAGCAGTGGGGGCCCCC AACTAGGAAGCATCATTAGACACTCAACCAACCACTTCTTGAAGCAGTGTCTACTCT -----

U
T
R
A
N
S
L
A
T
E
D
F
R
O
M
R
A
T
C
D
N
A
T
O
H
U
M
A
N
A
N
D
M
U
S
C
D
N
A

rat_cDNA human 5+3 corrected mus_cDNA_5	AGTGCCCACTCCTTCCCCACCTCTTTTACTAAGGGTGTGGTTACAGACAGCAAGTCAC AATGCCAGTTCOCATCTCCCTCCCTTTACTCAGAGAGCAGTTACTGACAACTGGGAC
rat_cDNA human 5+3 corrected mus_cDNA_5	ATCAGCTTTCCAGATGACGTCAAATAGAGTGGTCACCATATATGAATCTTCAAGGCACAA TOCCATTTCCGGGCTTATGACAAATACAGTGGTCAAGCTGCACGAATCCTCAAGGCACAA
rat_cDNA human 5+3 corrected mus_cDNA_5	TACAGATCTGCAGCAACCTCAGCAGAGGCTAGCCCCAATCCTGAGATCATAACTGGAAC TCNNNNNNNNCA--AATGCCAAGTTCA-----CINNAATTGNGAACNNNNNACTCNNNN
rat_cDNA human 5+3 corrected mus_cDNA_5	CACTGACTCTCCCTCTAATCTGTTTCCATCCACTTCTGTGCCAGCACTAAGGGTAGATAA NACTTCATCTACNTCTAATCTGTTACATTCTACTCCCATGCCAGCACTAACCAAGTTAA
rat_cDNA human 5+3 corrected mus_cDNA_5	ACCACAGAATTCTAAATGGAAGCCCTCTCCCTGGCCAGACACAAATATCAGCTCAAGTC ATCACAGAATTCCAAATTAAGTCCATCTCCCTGGGCGAATACCAATTTGGCACAACCC
rat_cDNA human 5+3 corrected mus_cDNA_5	ATACTCCGAAACCATTTAGAGAGGGCAAAAGGCCAGCAGTAAGCATGTCCCCCACCCTCAG ATACTCAGACATTGCTGAAAAGGCCAAAAGCCAGAAGTAAGCATGTTGGCT-ACTACAG
rat_cDNA human 5+3 corrected mus_cDNA_5	-CCTTCCAGAGGCCAGCACTCATGCCCTCACACTGGAATACACAGAAGCATGCAGAAAGA GCCGTGTCGAGGCCACCCTCTGTTTCAGATTGGGATGGACAGAGAACACAAAGAGA
rat_cDNA human 5+3 corrected mus_cDNA_5	GTGTTTTTGATAAGAAACCTGGTCRA-AACC--CAACTTCCAAACATCTGCCCTACGTC GTGACTTTGATAAGAAACAGTTCAAGAAGCAACCACTTCCAACTCCTTCCCTTTGACT
rat_cDNA human 5+3 corrected mus_cDNA_5	CTCTACCTAAGACTCTATTGAAAAAGCCAGAATAATTGGAGGAAAGGCTGCAGCTTTA CTTTGTCTAGGTATATATTGAAAAGCCAGGATAGTTGGAGGAAAAGCTGCAAGTTTA
rat_cDNA human 5+3 corrected mus_cDNA_5	CAGTTCAGCTAATTCAGACGTTTTCTTCTCTGAGGCTGTTGGAGACCACTGCCCA CTATTCCAGCTAACTCAGATGCCCTTCTTCCCTGTGAAGCTGTTGGAATCCCTGCCCA
rat_cDNA human 5+3 corrected mus_cDNA_5	TCATCCACTGGACCAGAGTTTCATCAGGANTTGAATATCCCAAGGGACACAGAAAAGCC CCATTTCATTGGACCAGAGTNNNNTCAGGACTTGATTATCTAAGAGGAAACAGAAATAGCA
rat_cDNA human 5+3 corrected mus_cDNA_5	GGTTCACGTCGCTTCCCAATGGCACCTTGTCCATCCAGAGGGTCAGTATTCAGGACCGTG GGGTCCAGGTTCTCCCCAATGGTACCTGTCCATCCAGAGGGTGGAATTCAGGACCGCG
rat_cDNA human 5+3 corrected mus_cDNA_5	GACAGTACCTGTGCTCTGCATTTAATCCACTGGGCGTAGACCATTTTCATGCTCTTTGT GACAGTACTTGTTGTTCCGCATCCCAATCTGTTTGGCAGACCACTTCATGTCACCTTGT
rat_cDNA human 5+3 corrected mus_cDNA_5	CTGTGGTTTTTTAOCOGGCAAGGATTTTGGACAGACATGTCAAGGAGATCACAGTTCACT CTGTGGTTTCTATCTCTCCAGGATCCTGAGAGACGTACCAAGAGATCACAGTTCACT

rat_cDNA human_5+3 corrected mus_cDNA_5	TTGGAAGTACTGTGGAACATAAGTGACAGTGGAGGGTATGCCGAGGCCTACGGTTTCCT COGGAAGCACTGTGGAACGAAGTGACAGGACAGAGGTAGGCCAAGCCCTACAGTTAOCCT -----
rat_cDNA human_5+3 corrected mus_cDNA_5	GGATACTTGCAAACCAAACGGTGGTCTCAGAAACGGCCAGGGAAGCAGAAAGGTCTGGG GGATTCTTGCAAACCAAACAGTTGTCTCAGAATCATCCAGGGAAGTAGGCAGGCTGTGG -----
rat_cDNA human_5+3 corrected mus_cDNA_5	TAAACCTGATGGAACATTGATCATCTATAATCTGAGTCTTTATGATCGTGGTTTTTACA TGACGGTTGACGGAACATTGGTCTCCACAATCTCAGTATTTATGACCGTGGCTTTTACA -----
rat_cDNA human_5+3 corrected mus_cDNA_5	AGTGTGTGGCCAGCAACCCATCTGGCCAGGATTCACTGTTGGTTAAGATACAAGTCATCA AATGTGTGGCCAGCAACCCAGGTGGCCAGGATTCACTGCTGGTTAAATACAGTCATTG -----
rat_cDNA human_5+3 corrected mus_cDNA_5	CAGCTCCOCCTGTCAATATAGAGCAAAAGAGGCCAAGCCATCGTTGGGGTTTTAGGTGGAA CAGCACCACTGTATTCTAGAGCAAAAGAGGCCAAGTCATTGTAGGCACCTTGGGGTGAAA -----
rat_cDNA human_5+3 corrected mus_cDNA_5	GTTTGAAACTGCCCTGCACTGCAAAAGGAACCTCCCGCCCTAGTGTTCCTGCTGGTCTTT GTTTAAACTGCCCTGTACTGCAAAAGGAACCTCCTCAGCCAGCGTTTACTGGGTCTCT -----
rat_cDNA human_5+3 corrected mus_cDNA_5	ATGATGGGACTGAACATAAAACCATTCAGTTGACTCATTCAGATTTTTCTGTATCCAA CTGATGGCACTGAAGTGAAACCATTCAGTTTACCAATTCAGTTGTTCTTATTTTCAA -----
rat_cDNA human_5+3 corrected mus_cDNA_5	ATGGAACTCTGTATATAAGAAGCATCGCTCCTTCAGTGGGGGCACCTTATGAGTGCAATG ATGGGACTTTGTATATAAGAAACCTAGCCTCTTCAGACAGGGGCACCTTATGAATGCATTG -----
rat_cDNA human_5+3 corrected mus_cDNA_5	CCACCACTCCTCAGGCTCAGAGAGAAGGGTAGTGATTCTTACTGTGGAAGAGGGAGAGA CTACCACTTCCACTGGTTTCGGAGCGAAGAGTAGTAATGCTTACAATGGAAGAGCGAGTGA -----
rat_cDNA human_5+3 corrected mus_cDNA_5	CAATCCCCAGGATAGAACTGCTCTCAGAAATGGACTGAGGTGAATTTGGGTGAGAAAT CCAGCCCCAGGATAGAACTGCTATCCAGAAAGGACTGAAGTGAATTTGGGGACAAAT -----
rat_cDNA human_5+3 corrected mus_cDNA_5	TACTACTGAACTGCTCAGCTACTGGGGATCCAAAGCCTAGAATAATCTGGAGGCTGCCAT TACTACTGAACTGCTCAGCCACTGGGGAGGCCAAACCCCAATAATGTGGAGGTTACCAT -----
rat_cDNA human_5+3 corrected mus_cDNA_5	CCAGGCTGTCTACGACCACTGGCCACAGAATGGGCAGCCGAATCCAGCTACCCAAATG CCAAGGCTGTGGTCGACCACTGG-----GCAGCTGGATCCAGCTCTACCCAAATG -----
rat_cDNA human_5+3 corrected mus_cDNA_5	GATCCTGTGGTGGTTGGGTGAGTGACGGAAAGACGCTGGTGACTACTTATGTGTGGCAA GATCCCTGTTTATTTGGATCAGTAACAGAAAGACAGTGGTGTCTACTTGTGTGGCAA -----
rat_cDNA human_5+3 corrected mus_cDNA_5	GAAACAAATGGGAGATGACCTAGTCTGATGATGTCCGCCTGAGATTGACACCTGCCA GAAACAAATGGGGGATGATCTGATACTGATGATGTAGCCTAAGACTGAAACCTGCCA -----

rat_cDNA human 5+3 corrected mus_cDNA_5	AAATTGAACAGAAAGCAGTATTTTAAAGCAAGTGCTCCATGGGAAAGATTTCAGTGTG AAATTGACCCACAGCAGTATTTTAGAAGCAAGTGCTCCATGGGAAAGATTTCAGTGTG -----
rat_cDNA human 5+3 corrected mus_cDNA_5	ACTGCAAGGCTCTGGCTCCCTGTGCTGAGGTATCTGGAGTTTGCTGATGGGACAG ATTGCAAGCTTCCGGCTCCCTAGTGCCAGAGATATCTTGGAGTTTGCTGATGGAAOCA -----
rat_cDNA human 5+3 corrected mus_cDNA_5	TGCTCAACAATGTAGCCCAAGCTGATGACAGTGGCTATAGGACCAAGAGGTACACCTTT TGATCAACAATGCAATGCAAGCCGATGACAGTGGCCACAGGACTAGGAGATATACCTTT -----
rat_cDNA human 5+3 corrected mus_cDNA_5	TCCACAATGGAACTTTGATTTCAACAAGCTTGGGATGGCAGAGGAAGGAGATTATATCT TCAACAATGGAACTTTTATCTTCAACAAGTTGGGATAGCGGAGGAAGGAGATTATACTT -----
rat_cDNA human 5+3 corrected mus_cDNA_5	GCTCTGCCAGAACACCTTAGGGAAGATGAGATGAAGTCCACCTAACAGTTCTAACAG GCTATGCCAGAACACCTTAGGGAAGATGAAATGAAGGTCCACTTAACAGTTATAACAG -----
rat_cDNA human 5+3 corrected mus_cDNA_5	CCATCCACGGATAAGGCAAGCTACAAGACCACCATGAGGCTCAGGGCTGGGAAACAG CTGCTCCCGGATAAGGCAGAGTAACAACCAACAGAGAAATCAAAGCTGGAGACACAG -----
rat_cDNA human 5+3 corrected mus_cDNA_5	CTGTCTTGACTGCGAGGTCACTGGGGAACCGAAGCCCAATGTATTTGGTTGCTGCCTT CTGTCTTGACTGTGAGGTCACTGGGATCCCAACCAAAATATTTGGTTGCTGCCTT -----
rat_cDNA human 5+3 corrected mus_cDNA_5	CCAACAATGTCAATTCATTCTCCAATGACAGGTTCAATTTATGCCAATAGACTTTGT CCAATGACATGATTTCTTCTCCATGATAGGTACACATTTATGCCAATGGGTCTTTGA -----
rat_cDNA human 5+3 corrected mus_cDNA_5	CCATCCATAAAGTGAAACCACTTGACTCTGGGACTATGTGTGCGTAGCTCAGAACTTA CCATCAACAAAGTGAAACTGCTCGAATCTGGAGAGTAAGTATGTGTAGCCCGAAATCCA -----
rat_cDNA human 5+3 corrected mus_cDNA_5	GTGGGATGACACTAAGACATACAACTGGACATTTGTCTTAAACCTCCATTATCAATG GTGGGATGACACCAAAATGTACAACTGGATGTGTCTTAAACCTCCATTATCAATG -----
rat_cDNA human 5+3 corrected mus_cDNA_5	GCCTGTATGCRAACAAGACTGTTATTAAGCCACAGCCATTGGGCACTCCAAAAATACT GTCTGTATACAAACAGAACTGTTATTAAGCCACAGCTGTGAGACATTCAAAAACACT -----
rat_cDNA human 5+3 corrected mus_cDNA_5	TTGACTGCAGAGCAGATGGGATCCCATCTTCCAGGTCACTGGATTATGCCAGGCAATA TTGACTGCAGAGCTGAAGGGACACCATCTCTGAAGTCATGTGGATCATGCCAGCAATA -----
rat_cDNA human 5+3 corrected mus_cDNA_5	TTTTCTCCAGCTCCATACTTTGGAGCAGAGTCACGGTCCATCCAAATGAACTTGG TTTTCTCAGAGCCCATACTATGGAAGCAGAAATCAGAGTCCATAAAATGAACTTGG -----
rat_cDNA human 5+3 corrected mus_cDNA_5	AGATGAGGAACATCCGGCTTCTGACTCTGCGGACTTCACTGTGTGGTTCGGAGCGAGG AAATTAGGAATGTGAGGCTTTCAGATTCAGCGACTTTATCTGTGTGGCCGAAATGAAG -----

rat_cDNA human 5+3 corrected mus_cDNA_5	GAGGAGAGAGTGTGTTGGTAGTGCAGTTAGAAGTCTAGAAATGCTGAGAAGACCAACAT GTGGAGAGAGCGTGTGGTAGTACAGTTAGAAGTACTGGAAATGCTGAGAAGACCGACAT -----
rat_cDNA human 5+3 corrected mus_cDNA_5	TCAGAAACCCATTCAACGAAAAAGTCATCGCCCAAGCTGGCAAGCCCGTAGCACTGAAC TTAGAAATCCATTTAATGAAAAATAGTTGCCAGCTGGGAAAGTCCACAGCATTGAATT -----
rat_cDNA human 5+3 corrected mus_cDNA_5	GCTCTGTGGATGGGAACCCCACTGAAATTAOCTGGATCTTACCTGAOGGCACACAGT GCTCTGTGATGGTAACCCACCACTGAAATAATCTGGATTTTACCAATGGCACACGAT -----
rat_cDNA human 5+3 corrected mus_cDNA_5	TTGCTAACAGACCAACAATTCCCGTATCTGATGGCAGGCAATGGCTCTCTCATCTTT TTTCCAAATGGACCAAAAGTTATCAGTATCTGATAGCAAGCAATGGTTCTTTTATCATT -----
rat_cDNA human 5+3 corrected mus_cDNA_5	ACAAAGCAACTCGGAACAAGTCAGGGAAGTATCGCTGTGCAGCCAGGAATAAGGTTGGCT CTAAACAACCTCGGGAGGATGCAGGAAAATATCGCTGTGCAGCTAGGAATAAAGTTGGCT -----
rat_cDNA human 5+3 corrected mus_cDNA_5	ACATCGAGAACTCATCTGTAGAGATTGGGCAGAAGCCAGTCATTCTGACATACGAAC ATATTGAGAAATTAGTCATATTAGAAATTGGCCAGAAGCCAGTTATTCTTACCTATGCAC -----
rat_cDNA human 5+3 corrected mus_cDNA_5	CAGGGATGGTGAAGAGCGTCAGTGGGGAAACGTTATCACTGCATTGTGTGTCTGATGGGA CAGGGACAGTAAAAGGCATCAGTGGAGAATCTCTATCACTGCATTGTGTGTCTGATGGAA -----
rat_cDNA human 5+3 corrected mus_cDNA_5	TCCCCAAGCCAAATGTCAAGTGGACTACACCGGGTGGCCATGTAATCGACAGGCGCTCAAG TCCCTAAGCCAAATATCAATGGACTATGCCAAGTGGTTATGTAGTAGACAGGCGCTCAAA -----
rat_cDNA human 5+3 corrected mus_cDNA_5	TGGATGGAAAATACATACTGCATGAAATGGCAGCTGGTCATCAAAGCAACACAGCTC TTAATGGGAATACATATTGCATGACATGGCACCTTAGTCATTAAAGAGCAACAGCTT -----
rat_cDNA human 5+3 corrected mus_cDNA_5	ACGACCAAGGAAATATATCTGTAGGGCTCAAAACAGTGTGGCCAGGCAGTTATTAGCG ATGACAGAGGAACTATATCTGTAGGGCTCAAAATAGTGTGGTCATACACTGATTACTG -----
rat_cDNA human 5+3 corrected mus_cDNA_5	TGTCAGTGATGGTTGTGGCTACCTCCCGAATCATAAACTAOCCTACCCAGGAACATGC TTCCAGTAATGATTGTAGCTACCTCCCGAATTACAAATGTCCACCCAGGAGTATTG -----
rat_cDNA human 5+3 corrected mus_cDNA_5	TCAGGAGGACAGGGGAAGCCATGCAGCTCCACTGTGTGGCTTGGGAATCCCCAAGCCAA TCACCAGGACAGGGGCGCCCTTTCAGCTCCACTGTGTGGCTTGGGAGTTCCCAAGCCAG -----
rat_cDNA human 5+3 corrected mus_cDNA_5	AAGTCACCTGGGAGCGCCAGACACTCCCTGCTCTCAAAGCAACAGCAGAAACCCC AAATCACATGGGAGATGCCTGACCACTCCCTCTCTCAACGGCAAGTAAGAGAGGACAC -----
rat_cDNA human 5+3 corrected mus_cDNA_5	ATAGAAGTGAGATGCTTACCCACAAGGTACGCTGGTCATTGAGAATCTCCAAACCTCGG ATGGAAGTGAGCAGCTTCACTTACAGGTACCTTAGTCATTGAGAATCCCAACCTCGG -----

rat_cDNA
human_5+3 corrected
mus_cDNA_5

ATTCCGGAGTCTATAAGTGCAGAGCTCAGAACCTACTGGGACTGATTACGCAACAACTT
ATTCTGGGATATACAAATGCACAGCAAGAAACCCACTTGGTAGTGATTATGCAGCAAGCT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

ACATCCAGGTACTCTGACAGGAGGGGGAGACTAAAATTCAACAGAGTCCACATCCACA
ATATTCAAGTAATCTGACATGAA-----ATAATAAGT-CAACAA-----CATCTGGGCA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

GGGTTTATTTTTTGGAGAAGTTTAATCAAAGGCAGCCATAGGCATGTAAATGAGTCTGA
GAATTTATTTTTTGGAGAAGTTTAATCAAAGGCAGCCATAGGCATGTAAATGAATTTGA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

ATACATTACAGTATTAAATTTACAATGGACATGCGA--TGA---GACTTGTAATGAAA
ATACATTACAGTATTAAATTTACAATGAACATGCAAAATAAAAGGACTTGTAATTAAT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

GCATTGTGAAGTGA-----ACCGAGTCTCTG--TGGATCTCAAAGCAAACCTTTAACTTAA
GCATTATGAAGTGAATGATGACTGATTTTAAATGGATCTCAAACAAACCTTTAACTTAA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

GGCACTTTGATTTTTGCCAACAAATAAACAACATTAAAGAGAAAAAATGATCCACTAC
GGCACTTTTATTTTTGCCAACAAATAAACAATAACAA-----ACATTGAAACGGTTCACTAT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

GAAATAACAAACGGCTAATGCACCTGAATTCT-CAGTAAAAAGACCTTTCTCTCGCTAAC
AAAATAACAAATGGCTAATGTACCTGAATTTTTCAGTAAAAAAA--TGAACCTT-CTAAT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

AGTTGCCAGCTGCCTCGTGTCTGTTTCTACCAATGTCAAAACATCGCACACAGGGTGA
A-----CCAGTTGCCTAGTGTCCACCTCTATCAATGTTACAAGCATGGCACTCAG-----A

rat_cDNA
human_5+3 corrected
mus_cDNA_5

ATGGAGTCAACGGGAAGATTAGTTTGGGGTCTGTGTAATCTCAATGTACAAATATTC
ACAGAGCAATGGAAAAATATTAAATCTGC-----AATCT--ATGTATAAATATTT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

TGTCNCTGGTTTATAAATTTTT-GATAAAACCGAAAAAATAAAAAAAAAAAAAAAAAA
TGT-----GGTTTATAAATTTTTTGTCTAAACCTACAGAAATAG-----

rat_cDNA
human_5+3 corrected
mus_cDNA_5

AAAA

(rat_cDNA: SEQ ID NO:7)
(human_5+3 corrected: SEQ ID NO:8)
(mus_cDNA_5: SEQ ID NO:9)

Figure 15

rat	MQVRGREVSGLLISLTAVCLVVTTPGSRACPRRCACYPTEVHCTFRYLTSIPDGI PANVE
human_5+3_corrected	MRVKGRGITCLLVSPFAVICLVATPGGKACPRRCACYPTEVHCTFRYLTSIPDSIPP NVE
mouse_5_corrected	MOKRGREVSCLLISLTATCLVVTTPGSRVCPRRACACYPTEVHCTFRDLTSIPD-GPANVE
rat	RINLGYNLSLTRLTENDEDFGLSKLEIMLHLSNGIHRVSDRTTSGLSLQVLKMSYNKVQII
human_5+3_corrected	RINLGYNLSLRLMETDFSGLTKEIMLHLSNGIHTIPDKTTSDLQALQVLKMSYNKVRKL
mouse_5_corrected	RVNLGYNSLTRLTENDEFSGLSKLEIMLHLSNGIHRVSDRTTSGLSLQVLKMSYNKVQII
rat	RKDTFYGLGSLVRLHLDHNNIEFINPEAFYGLTSLRLVHLEGNRLTKLHPDTFVSLSYLQ
human_5+3_corrected	QKDTFYGLRSLTRLHMDHNNIEFINPEVFGINFLRLVHLEGNQLTKLHPDTFVSLSYLQ
mouse_5_corrected	EKDTLYGLRSLTRLHLDHNNIEFINPEAFYGLTLLRLVHLEGNRLTKLHPDTFVSLSYLQ
rat	IFKTSFIKYLELSDNFLTSLPKEMVSYMPNLESYLHGNPWTCDCHLKNLSEWMQGNPDI
human_5+3_corrected	IFKISFIKFLYLSDNFLTSLPQEMVSYMPDLDSLYLHGNPWTCDCHLKNLSDWIQ--PDV
mouse_5_corrected	IFKTSFIKLYLYDNF-TSLPKEMVSSMPNLESYLHGNPWTCDCHLKNLSEWMQGNP--
rat	IKCKKDRSSSSPQQCPLOMNPRI SKGRPFAMVPSGAFLCTKPTIDPSLKSLSLVTQEDNG
human_5+3_corrected	IKCKKDRSPSSAQQCPLCMNPRTSKGKPLAMVSAAFQCAKPTIDS6LKSLSLTILEDSS
mouse_5_corrected	-----
rat	SASTSPQDFIEFFGSLSLNMTXSGNKADMVCSIQKPSRTSPATTEENDYIMLNASFST
human_5+3_corrected	SAFISPPQGFMAFFGSLTINMTDQSGNEANMVCSIQKPSRTSPATTEENDYIVLNTSFST
mouse_5_corrected	-----
rat	NLVCSVDYNIHQPVWQLLALYSDSPLILERKPOLTETPSLSSRYKQVALRPEDIFTSIEA
human_5+3_corrected	FLVCNIDYGHIQPVWQIILALYSDSPLILERSHLLSETPOLYYKYKQVAFKPEDIFTNIEA
mouse_5_corrected	-----
rat	DVRADPFWFQQEKIVLQINRTATTLSTLQIQFSTDAQIALPRAEMRAERLKWIMIMNN
human_5+3_corrected	DLRADPSWIMQDQISLQINRTATTFSTLQIQYSSDAQITLPRAEMLPVKHKWTMISRDNN
mouse_5_corrected	-----
rat	PKLERTVLVGSTIALSCPGKGDPSFLENLLADGSKVRAPYVSEDGRILLDKNGKLELOM
human_5+3_corrected	TKLEHTVLVGSTVGLNCPGGQDTPHVDWLLADGSKVRAPYVSEDGRILLDKSGKLELOM
mouse_5_corrected	-----
rat	ADSF DAGLYHCISTNDADADVLTYRITVVEPYGESTHDSGVQHTVVTGETL DLPCLSTGV
human_5+3_corrected	ADSFDTGVYHCISSNYDDADILTYRITVVEPLVEAYQENGIIHTVFIGETL DLPCHSTGI
mouse_5_corrected	-----
rat	PDASISWILPGNTVFSQPSRDQIILNNGTLRILQVTPKQGHYQCVAA NPSGADFSSFKV
human_5+3_corrected	PDASISWVTPGNVLYQSSRDQKVLNNGTLRILQVTPKQGGYRCVAA NPSGVDFLI FQV
mouse_5_corrected	-----
rat	SVQKKQRMVEHDREAGGSGLGEYNSSVSLKQPASLKL SASALTGSEAGQVSGVHRKNK
human_5+3_corrected	SVKMKQRPLEHDGETEGSGLDESNIPIAHLEPPGAQLRTSALMEAEVGHKTSST6KRNN
mouse_5_corrected	-----
rat	HRDLIHRRGDSTLRFRFHRRLPLSARRIDPQRWAALLEKAKKNSVPKKQENTTVKPV
human_5+3_corrected	YRELTLORRGDSTHRRFRFENRRHFPPSARRIDPQHWAAALLEKAKKNAMPDKRENTTVSPP
mouse_5_corrected	-----


```

rat      WPEHKYQLKSYSETIERGKRPVSMSPHLSPLEASTHASHWNTQKHAESYVDKKPGQNP
human_5+3_corrected  WASYQFWHKPYSDLAEGKGPVSM LATTGLSEATTLLVSDWDGQKNTKSDFDKKPVQEA
mouse_5_corrected    -----

```

```

rat      RCAARNKVGYTEKLILLEIGQKPVILTYEPGVKSVSGEPLSLHCVSDGIPKPNVKWTFP
human_5+3_corrected RCAARNKVGYTEKLIVLEIGQKPVILTYAPGTVGIGBSLSLHCVSDGIPKPNIKWTFP
mouse_5_corrected

```

rat GGHVIDRPQVDGKYILHENGTLVIKATTAHDQGNVICRAQNSVGQAVISVSVMVVAIRK
human_5+3_corrected SGYVVDPRQINGKYILHDNGTLVIKEATAYDRGNICKAONSVGETLITVPVMIVAYPPR
mouse_5_corrected -----

rat IINYLPRNMLRRTGEANOLHCVALGIPKPKVTWETPRHSLLSKATARKPHRSEMLHPQGT
human_5+3_corrected ITNRPPRSIVTRTGAAFQLHCVALGVFKPEITWEMPDHSLSTASKERTHGSEQLHLQGT
mouse_5_corrected -----

rat LVIQNLQTSDSGVYKCAQNLGTDYATTYIQV
human_5+3_corrected LVIQNPQTSDSGIYKCTAKNPLGSDYAATTYIQV
mouse_5_corrected -----

(rat: SEQ ID NO:10)

(human_5+3_corrected: SEQ ID NO:11)

(mouse_5_corrected: SEQ ID NO:12)

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

[illegible]

```

rat      PITGVSRPSSSDISHTTADPSFSSHPSGSHTTASSLFHPRNNNTGNFPLSRHLGRERT
human_5+3_corrected SVREVSEPRNHFFYSHTTQILSTSTFPSPHTAAHSQFFIPRNS-TVNIPLFRRFRGRQK
      .! ***. * ..: ***** * *,*...***: * * * ****. * *: * *:***:

rat      IWSRGRVKNPHRTFVLRHRHRTVRPAIKGFANKNVBQVPATEYPMCHTCPSAEGLTVA
human_5+3_corrected  IGGRGRIISPRTFVLRHRHYSIFRSTTRGSSEKSTAFSATVLNVCLSCPLRERLTTA
      * ,***: .*:*****; .*: *:***:..! ..** * : * * **

rat      TAALSVPSSSHSALPKTNVGVIAEESTTVVKPILLFKDKQNVDEIITTTTKYSGGES
human_5+3_corrected  TAALSFPSPAAPITFFKADIARVPSEESTTLVQNPLLLLENKP--SVEKTPTIKYFRTEI
      *****,: :***: . * ,*****:***:***:; * .: * * * *

rat      NHVIPTASMTSAPTSVSLGKSPVDNSGHLSPGTIQTGKDSVETTPLPSPSTP--SIP
human_5+3_corrected  SQVTPTGAVMTAPTSIPMEKTHKVNASYPRVSTNEAKRDSVITSSLSGATKSPMTII
      .: * * * * ****; .: * :*.: .: * : :*** *: *...! * : *

rat      TSTKFSKRKTPLHQIFVNNQKKEGMLKNPYQFGLQKNPAKLPKIAPLLTGQSSPSDST
human_5+3_corrected  AITRFSRRKIPWQNFVNNENPKGRLRNQHKVSLQKSTAVMLPPTSALPQRQSSFFHFT
      ; *:***: * * : * ****; : * *: * !, ,***, * , *** : * * **** . *

rat      TLLTSPPPALSTTMAATQNKGTVEVVGARLSAGKKQPFNTSSPVLPTISKRSNLTNFI
human_5+3_corrected  TLSTSVMQIPSNLTITTAHTTTTKTHNPG-SLPTKKELPFPPLNPMPLPSIISKDSSTKSI
      ** ** * .*:***:!!!, *:. . . **.: *: ** .*:*** ** * . * !!

rat      STETPT-VTSPTATASVIMSETQRTSRKEAKDQING-P-RKNRNNANTTPROVSGYSAYS
human_5+3_corrected  STQTAIPATPTTFPASVITYETQTERSRAQTIQREQEPQKNRTPDNISPDQSSFTTPT
      **: * . ,*: * **** ** * *: . * : * ,***: * : * * **: :

rat      ALITADTFLAFSHSPRQDGGNVSAVAYHSTTS--LIATELFKKTOTLGNLTAEETL
human_5+3_corrected  AMTP--PALAFTHSPFENTTGISSTISFHSRTLNLTDVIEELAQASTQTLKSTIASETTL
      *: * . ,***:*** : * * !:*** * . * * : **** * * ****

rat      LSKSOESTTVKRAS-DTP-FLLSSGAPPVPTSPPPPFTKGVVTDKSVTSAFQMTSNRVV
human_5+3_corrected  SSKSHQSTTTTKASLDTPPPFLSSSATLMPVPISPPFTQRAVTDTRGDSRFLMTNTVV
      ***:***:; ** * * * ;***. . : * . ****: .***: : * !: ; * *

rat      TIYESSRHNTDLOQPSAEASNPBLITGTDSFNLFPSTVPEALKVLPKQNSKNKPSFW
human_5+3_corrected  KLEHSSRN--LQMPSSQLEP-----LTSSTSNLLHSTPMPLTTVKSONSKLTPSPFW
      .:***** ** *: : . * ,***: ** :*** . * **** ,****

rat      PEHKYQLKSYSETIEKGRPAVSMSPHLSLPEASTHSHWNTQKHAESVFDKPKGPQNP-
human_5+3_corrected  AEYQFWHKPYSDIAEKGRKEVSMLATTLGLSEATTLVSDWDGQKNTKSDPKKPVQEAT
      .*: : * .*: ****: * * * , .*.**.*. ,*: :*** ***** *:

rat      TSKHLPYVSLPKTLKKPRIIGGKAASFTVPANSDFLPCAVGDPPLPIIHWTVSSGXE
human_5+3_corrected  TSKLLPFDLSRYIFEKPRIVGGKAASFTIPANSDAFLPCAVGNPLETIHWTVB-GLD
      *** *: * ,: :***:*****;*****.*****;*** ***** * :

rat      ISQGTQKSRFHVLPNGTSLQIRVSIQDRGQYLCSAFNPLGVDHFHVSLSVVEYPARILDR
human_5+3_corrected  LSRGNQNSRVQVLPNGTSLQIRVEIQDRGQYLCSASNLFQTDHLHVTLSVVSYPPRILER
      :*: * .*: ;*****;*****.***** * : * .*: ;*** ** ,***: *

rat      HVKEITVHFGSTVELKCRVEGMPRTVSWILANQTVVSETAKGSRKVVTPDGTLLIYNL
human_5+3_corrected  RTKEITVHSGSTVELKRAEGRPSPTVTWILANQTVVSESSQGSROAVVTVDGTVLHNL
      :.***** *****.* * *:*****;***: . ** *****:***

rat      SLYDRGFYKCVASNPGQDLSLVKIQVITAPPVILEQKQAIQVGLGSLKLPCTAKGTP
human_5+3_corrected  SIYDRGFYKCVASNPGQDLSLVKIQVIAAPPVILEQRRQVIVGTWGESLKLPCAKGTP
      * ;*****;*****.*****:***:***. * *****

rat      QPSVHVVLVDGTTELKPLQLTHSRFLYPNGTLYIRSIAPSVRGTYECIATSSSGSERRVV
human_5+3_corrected  QPSVYVVLSDGTEVKPLQFTNSKFLFSNGTLYIRNLASSDRGTYECIATSSSGSERRVV
      ***:*** *****:***: * .*: ,*****: * , * *****;*****

rat      ILTVEEGETIPRIETASQKNTVENLGEKLLNCSATGDKPKRIIWRLPKAVIDQWHRMG
human_5+3_corrected  MLTMEERVTSPRIEASQKRTVENFGDKLLNCSATGEPKQIMWRLPKAVVDQ----G
      :*: * * ****:*** *****:***:*****;***: * :*****: * *

```

(rat: SEQ ID NO:13)
(human 5+3 corrected: SEQ ID NO:14)

Figure 17

MQKRGREVSCLLISLTAICLVVTPGSRVCPRRCACYPTEVHCTFRDLTSIPDGPANVER
VNLGYNSLTRLTENDFSGLSRLELLMLHSNGIHRVSDKTFSGLQSLQVLKMSYNKVQIIE
KDTLYGLRSLTRLHLDHNNIEFINPEAFYGLTLLRLVHLEGNRLTKLHPDTFVSLSYLQIF
KTSFIKXLYLYDNFTSLPKEMVSSMPNLESYLHGPNWTCDCHLKWLSEWMQGNP
(SEQ ID NO: 15)

Figure 18

MKVKGRGITCLLVSFAVICLVATPGGKACPRRCACYMPTEVHCTFRYLTSIPDSIPPVNE
RINLGYNLVRMLMETDFSGLTKELELLMLHSNGIHTIPDKTFSDLQALQVLKMSYNKVRK
LQKDTFYGLRSLTRLHMDHNNIEFINPEVFYGLNFLRLVHLEGNQLTKLHPDTFVSLSYL
QIFKISFIKFLYLSDNFLTSLPQEMSYPDLDSLLYLHGPNWTCDCHLRWLSDWIQPDVI
KCKKDRSPSSAQQCPLCMNPRTSKGKPLAMVSAAAFQCAKPTIDSSLKSKSLTILEDSSS
AFISPQGFMAPFGSLTLNMTDQSGNEANMVCSIQKPSRTSPIAFTEENDYIVLNTSFSTFL
VCNIDYGHIQPVWQILALYSDSPLILERSHLLSETPQLYYKYKQVAPKPEDIFTNIEADLR
ADPSWLMQDQISLQLNRTATTFTSLQIQYSSDAQITLPRAEVRPVKHKWTMISRDNNTK
LEHTVLVGGTVGLNCPGQGDPTPHVDWLLADGSKVRAPYVSEDGRILIDKSGKLELQM
ADSFDTG VYHCISSNYDDADILTYRITVVEPLVEAYQENGIIHTVFIGETLDLPCHSTGIP
DASISWVIPGNNVLYQSSRDKKVLNNGTLRLQVTPKDQGYRCVAANPSGVDFLIFQV
SVKMKGQRPLEHDGETEGSGLDESNP IAHLEKPPGAQLRTSALMEAEV GKHTSSTSKRH
NYRELTLQRRGDSTHRRFRENRRHFPPSARRIDPQHWAAALLEKAKKNAMPDKRENTTV
SPPPVVTQLPNIPGEEDDSSGMLALHEEFMVPATKALNLPARTVTADSRTISDSPMTNIN
YGTEFSPVVNSQILPPEPTDFKLSTAILTTAMSKNINPTMSSQIQGTTNQHSSSTVFPLLLG
ATEFQDSDQMGRGREHFQSRPPITVRTMIKDVNVKMLSSTTNKLLLESVNSHQTSVREV
SEPRHNHFYSHTTQILSTSTFPSDPHTAAHSQFPIPRNSTVNIPLFRRFGRQRKIGGRGRIIS
PYRTPVLRHRYSIFRSTTRGSSEKSTTAFSATVLNVTCLSCLPRELTTATAALSFPSAAPI
TFPKADIARVPSEESTTLVQNPLLLLENKPSVEKTTPTIKYFRTEISQVTPTGAVMTYAPT
SIPMEKTHKVNASYPRVSSSTNEAKRDSVITSSLGAIKPPMTIIAITRFSRRKIPWQQNFV
NNHNPKGRLRNQHKVSLQKSTAVMLPKTSPALPQRQSSPFHFRRLSTSVMQIPSNLTLT
AHHTTTKTHNPGSLPTKKELPFPLNPMLPSIISKDSSTKSIISTQTAIPATTPTFPASVITYE
TQTERSRAQTIQREQEPQKKNRTDPNISPQSSGFTTPTAMTPPALAFTHSPPENTTGISST
ISFHSRTLNLTDVIEELAQASTQTLKSTIASSETLSSKSHQSTTTTRKASLDTPIPPFLSSSAT
LMPVPISPPFTQRAVTDTRGDSHFRLMTNTVVKLHESSRHLNQMPSQLEPLTSSTSNLL
HSTPMPALTTVKSQNSKLTSPSWAEQFWHKPYSDIAEKGKKPEVSMLATTGLSEATTLV
SDWDGQKNTKKSDFDKPKPVQEATTSKLLPFDSLRYIFEKPRIVGGKAASFTIPANSDAF
LPCEAVGNPLPTIHWTRVSGLDLSRGNQNSRVQVLPNGTLSIQRVEIQDRFQYLCASNL
FGTDHLHVTLVSVSYPPRILERRTKEITVHSGSTVELKCRAGEGRPSPTVTWILANQTVVSE
SSQGSRQAVVTVDGTLVLHNSIYDRGFYKCVASNPGGQDSLLVKIQVIAAPPVILEQRR
QVIVGTWGESLKLPCAKGTPQPSVYWVLSDGTEVKPLQFTNSKLFLFSNGTLYIRNLAS

SDRFTYECIATSSTGSERRVVMLTMEERVTSPIREAASQKRTEVNFQDKLLLNCSATGEP
 KPQIMRLPSKAVVDQGSWIHYPNGSLFIGVTEKDSGVYLCVARNKMGDDLILMHVSLR
 LKPAKIDHKQYFRKQVLHGKDFQVDCKASGSPVPEISWSLPDGTMINNAMQADDSGHR
 TRRYRLFNNGTLYFNKVGVAEEGDYTCYAQNLTGKDEMKVHLTVITAAPRIRQSNKTN
 KRIKAGDTAAVLDCEVTGDPKPKIFWLLPSNDMISFSIDRYTFHANGSLTINKVKLLDSG
 EYVCVARNPSGDDTKMYKLDVVS KPPLINGLYTNRTVIKATAVRHKKHFDCAEGTP
 SPEVMWIMPDNIFLTAPYYGSRITVHKNGTLEIRNVRLSADFICVARNEGGEVLVVQLE
 VLEMLRRPTFRNPRNPFNEKIVAQLGKSTALNCVSDGNPPPEIHWILPNGTRFSNGPQSY
 QYLIASNGSFIISKTTREDAGKYRCAARNKVGYYIEKL VILEIGQKPVILTYAPGTVKGISGE
 SLSLHCVSDGIPKPNIKWTMPSGYVVD RPQINGKYILHDNGTLVIKEATA YDRGNYICKA
 QNSVGHTLITVPVMIVAYPPRITNRPPRSIVTRTGAA FQLHCV ALGV PKPEITWEMPDHS
 LLSTASKERTHGSEQLHLQGT LVIQNPQTSDSGIYKCTAKNPLGSDYAATYIQVI (SEQ ID
 NO: 16)

Figure 19

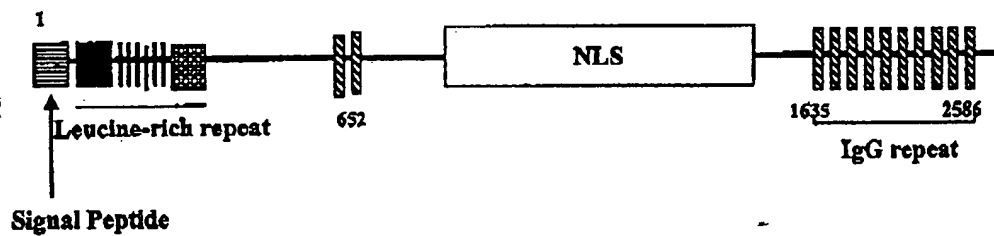


Figure 20

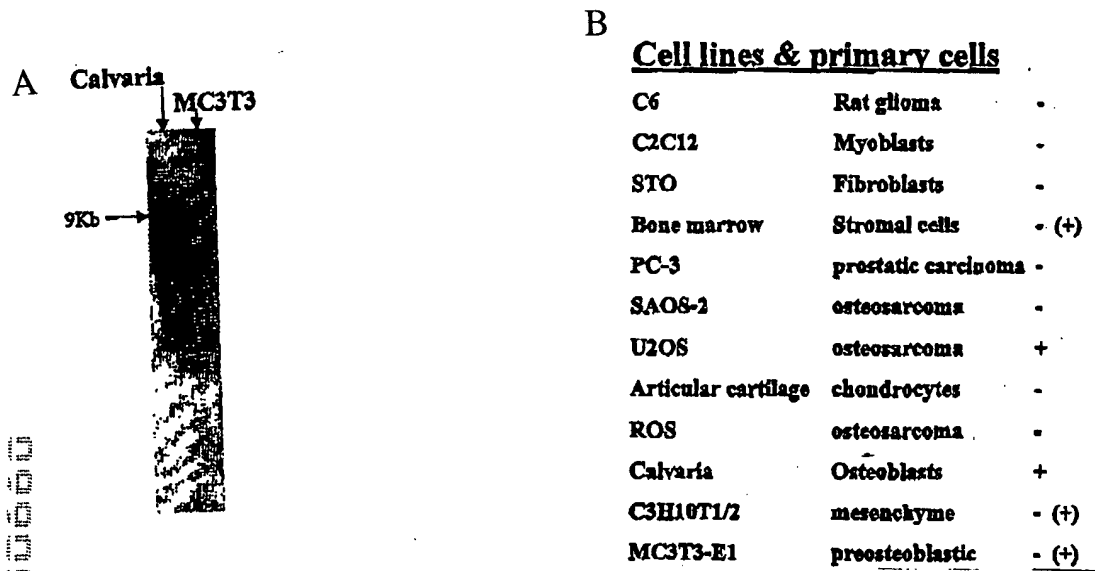


Figure 21

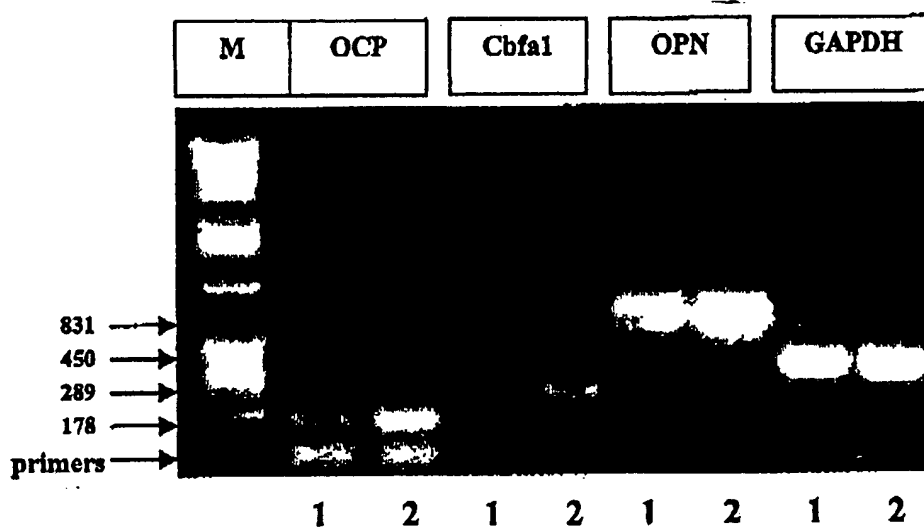


Figure 24

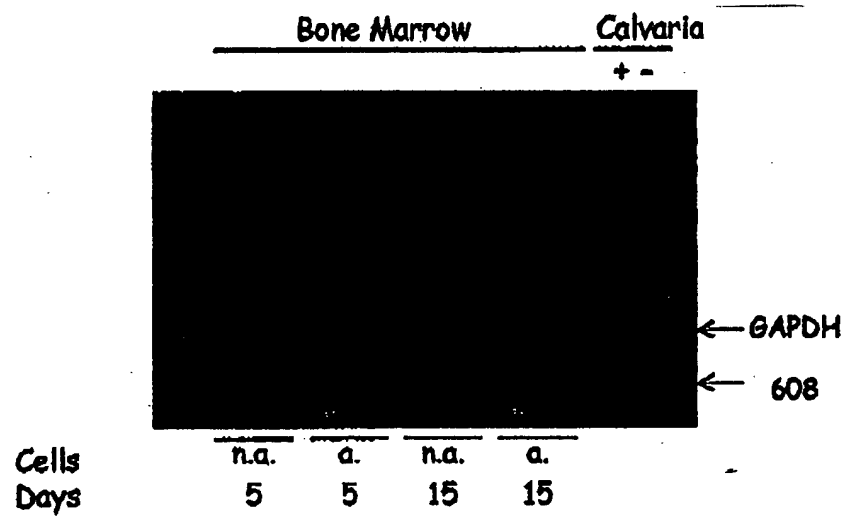
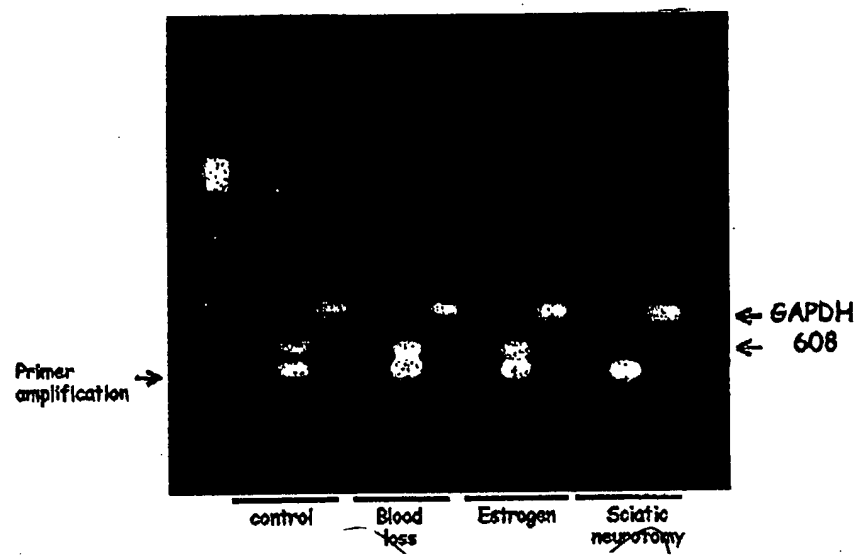


Figure 25



	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523</
--	---	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-------



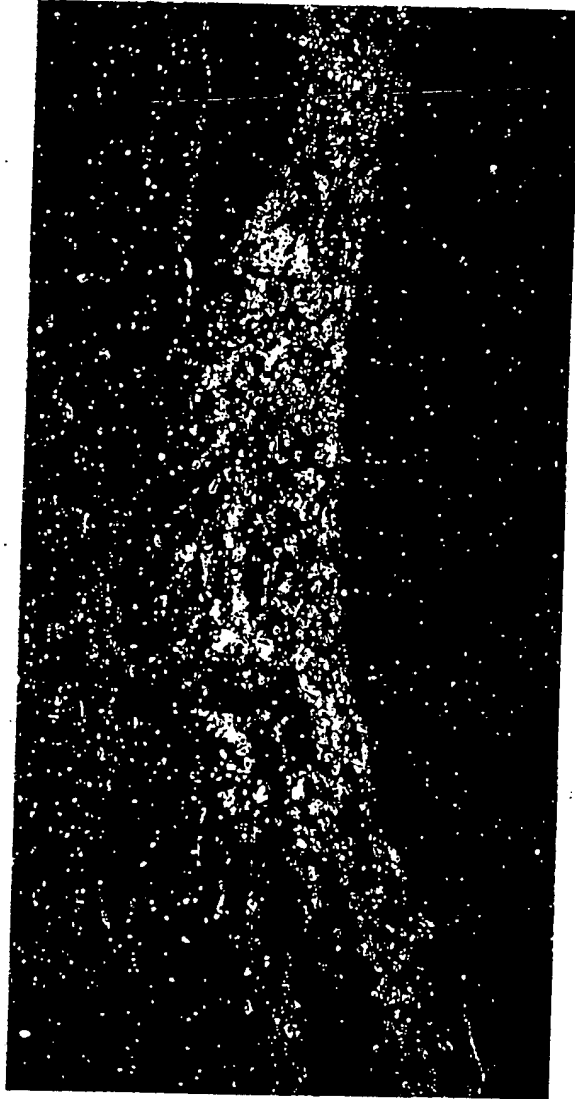
[illegible]

A high-contrast, black and white photograph showing a close-up of a textured surface, possibly a rock or concrete. The left side is bright and shows a rough, granular texture. The right side is dark and features a large, irregular, light-colored shape that appears to be a crack or a hole. The overall image has a grainy, high-contrast quality.



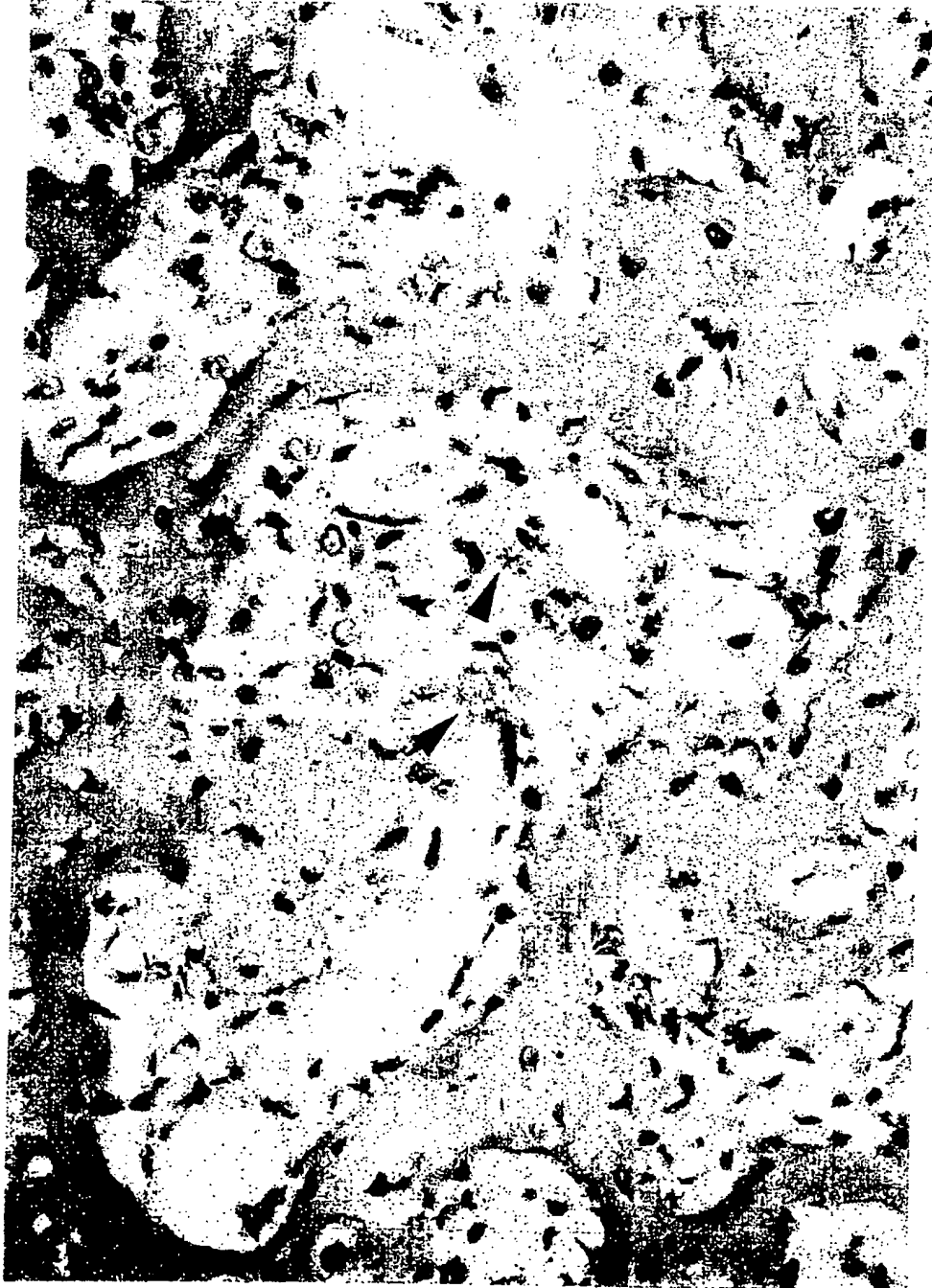
1

B

[illegible]

57/85

Figure 30



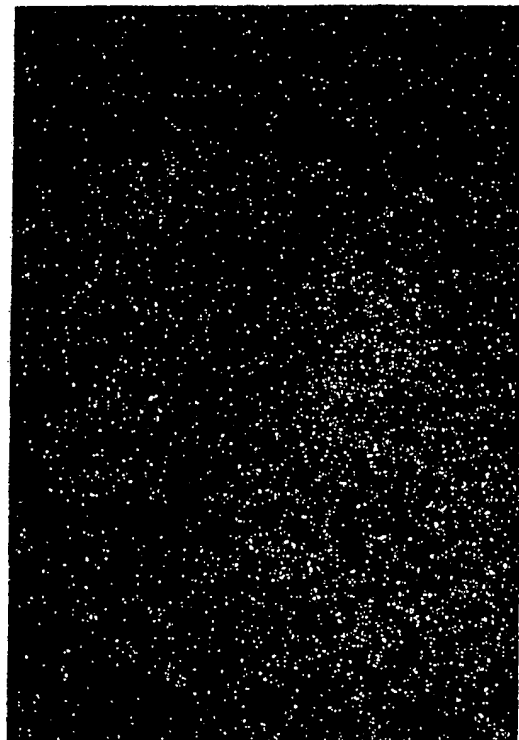
00005430 074304

[illegible]

A



B





This image is a high-contrast, black and white scan of a textured surface. It features a dense distribution of small, dark, irregular specks and fragments against a light, grainy background. The overall appearance is noisy and abstract, with no discernible text or geometric shapes. The dark elements vary in size and shape, some appearing as thin, elongated streaks while others are more rounded or irregular. The background is not uniform, showing a fine, granular texture.

61/85

Figure 34

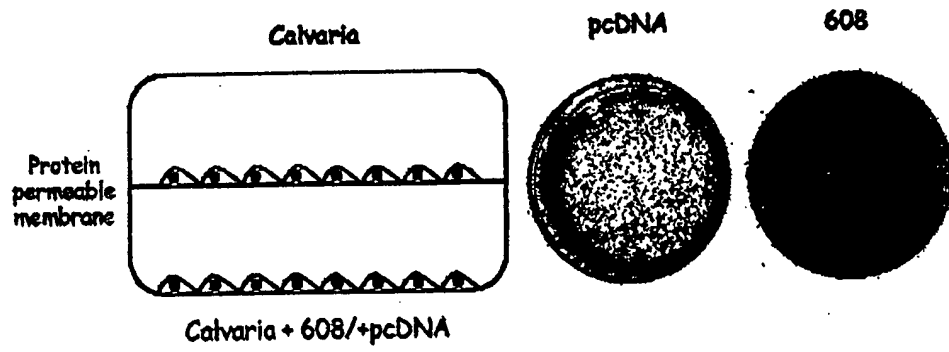
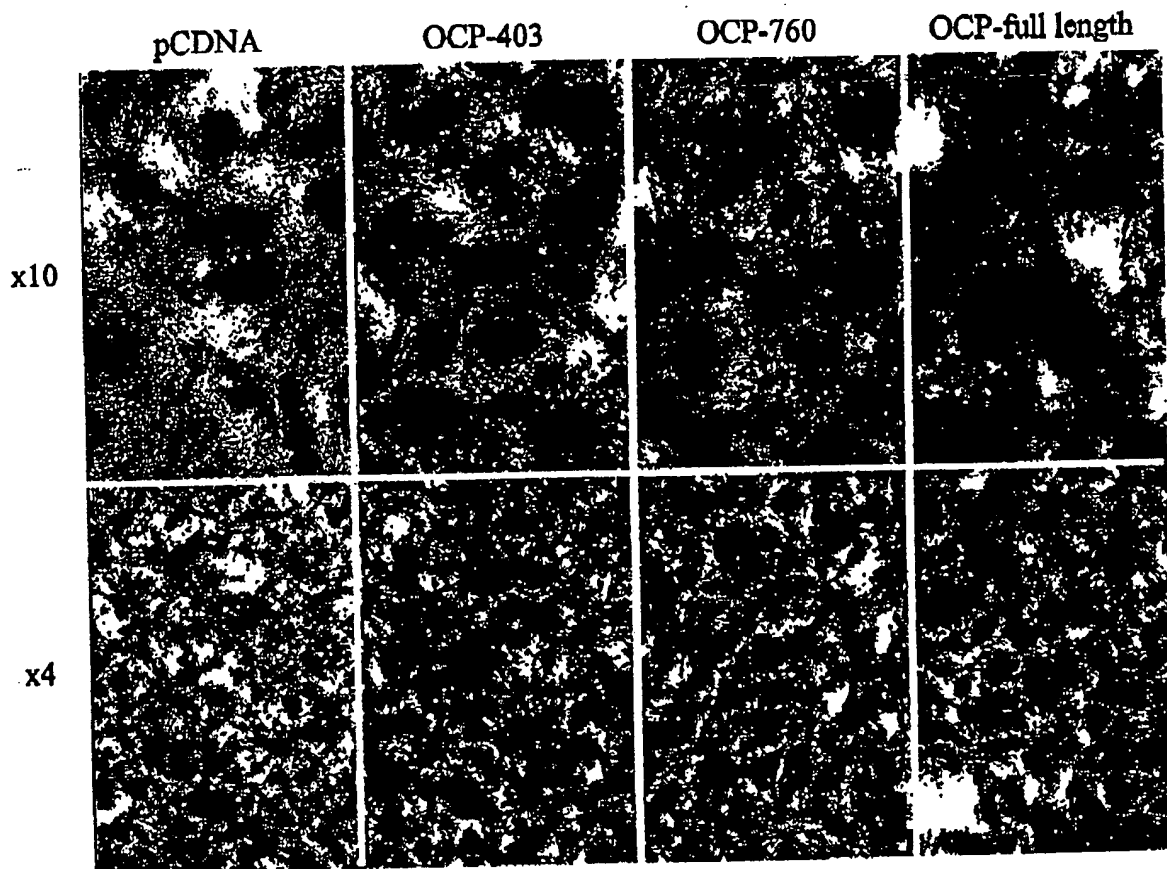


Figure 35



62/85

Figure 36

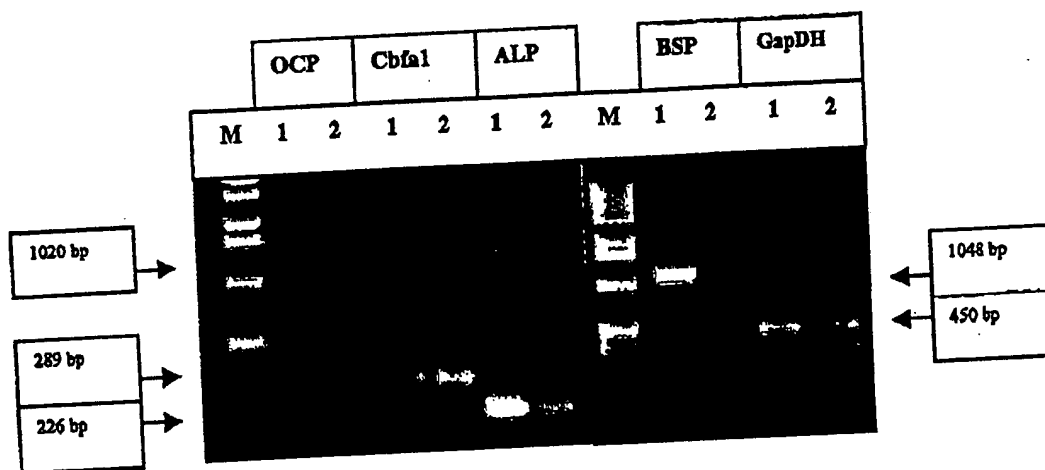


Figure 37

x4

x10

pCDNA
ROS stable line

OCP
ROS stable line



63/85

Figure 38

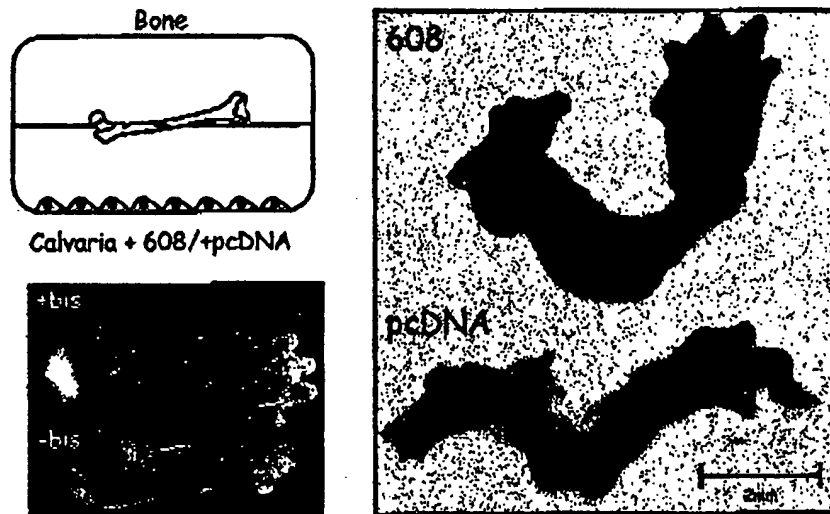
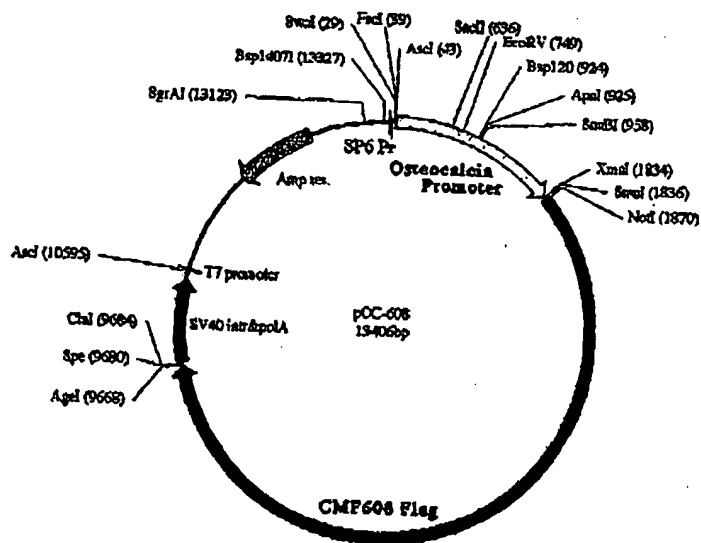
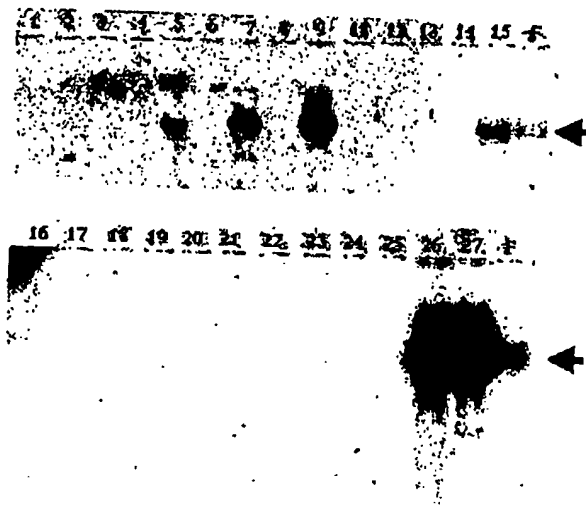


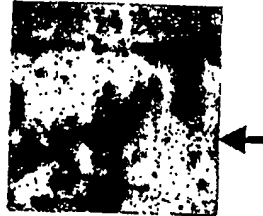
Figure 39



A



B



C

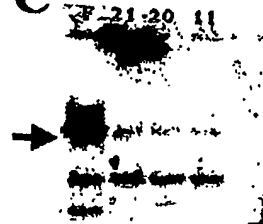


Figure 41

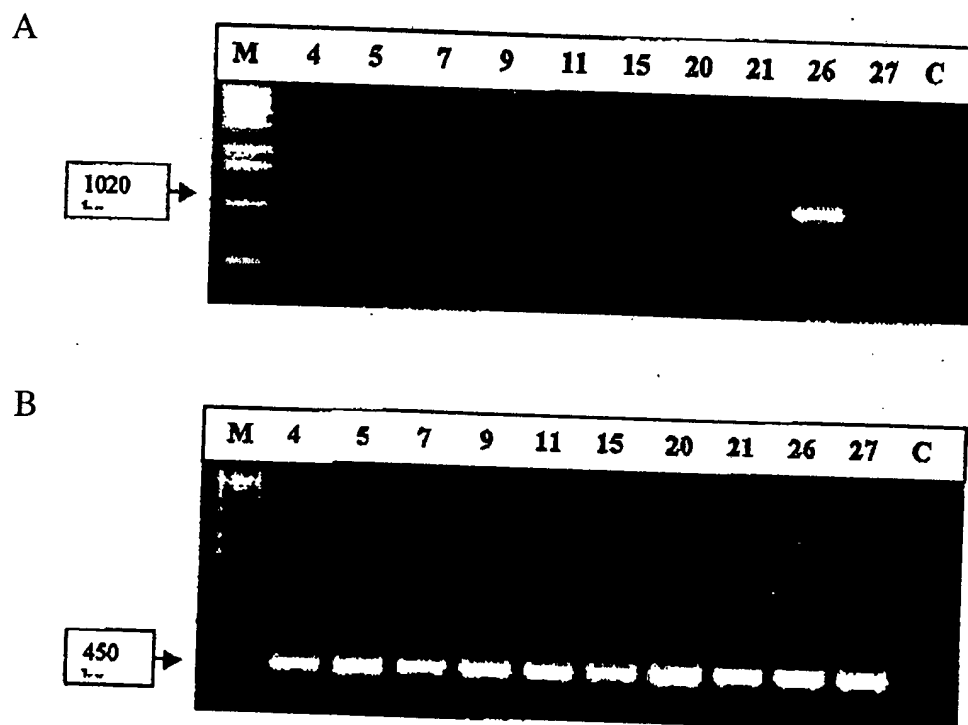
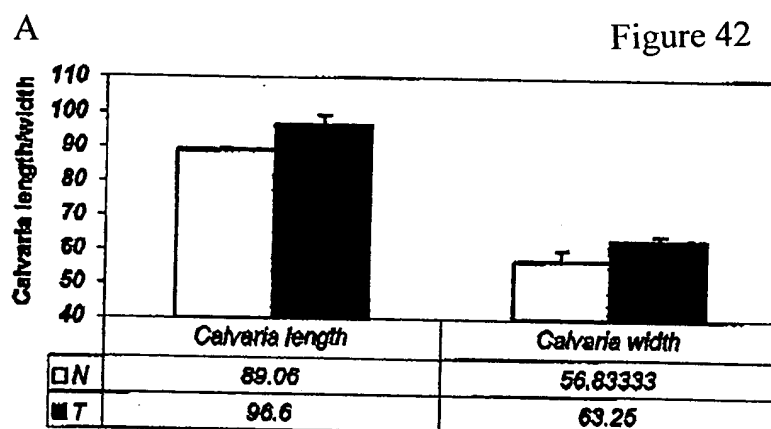


Figure 42



B

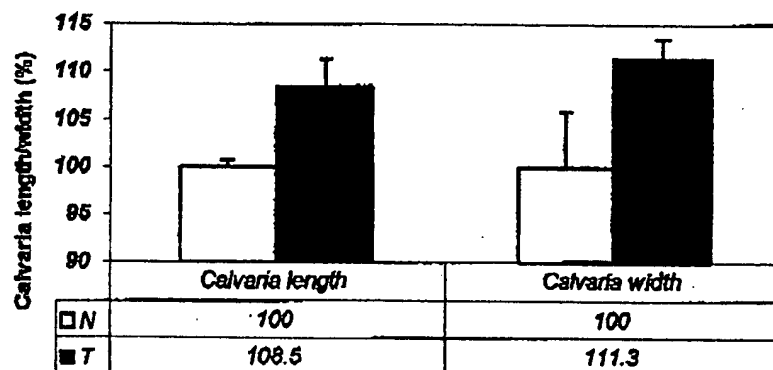


figure 43

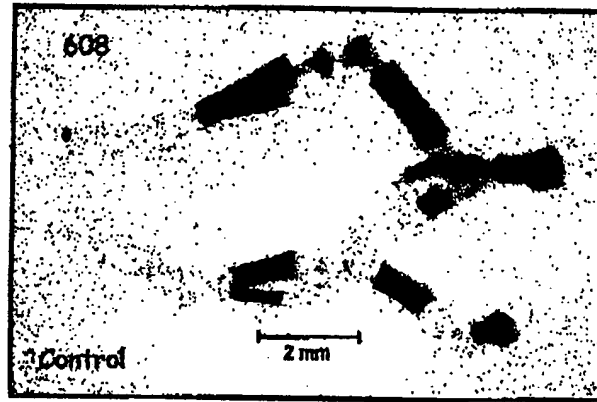


Figure 44

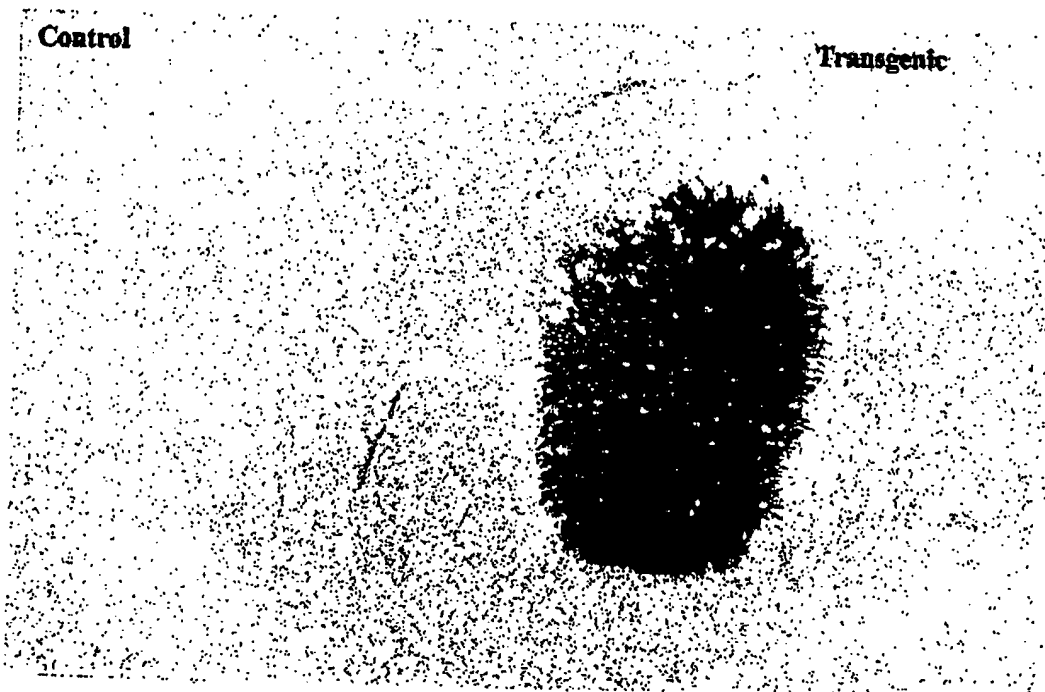
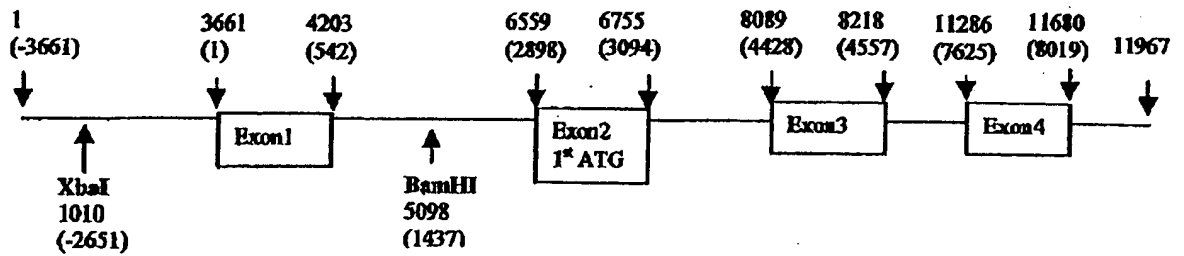


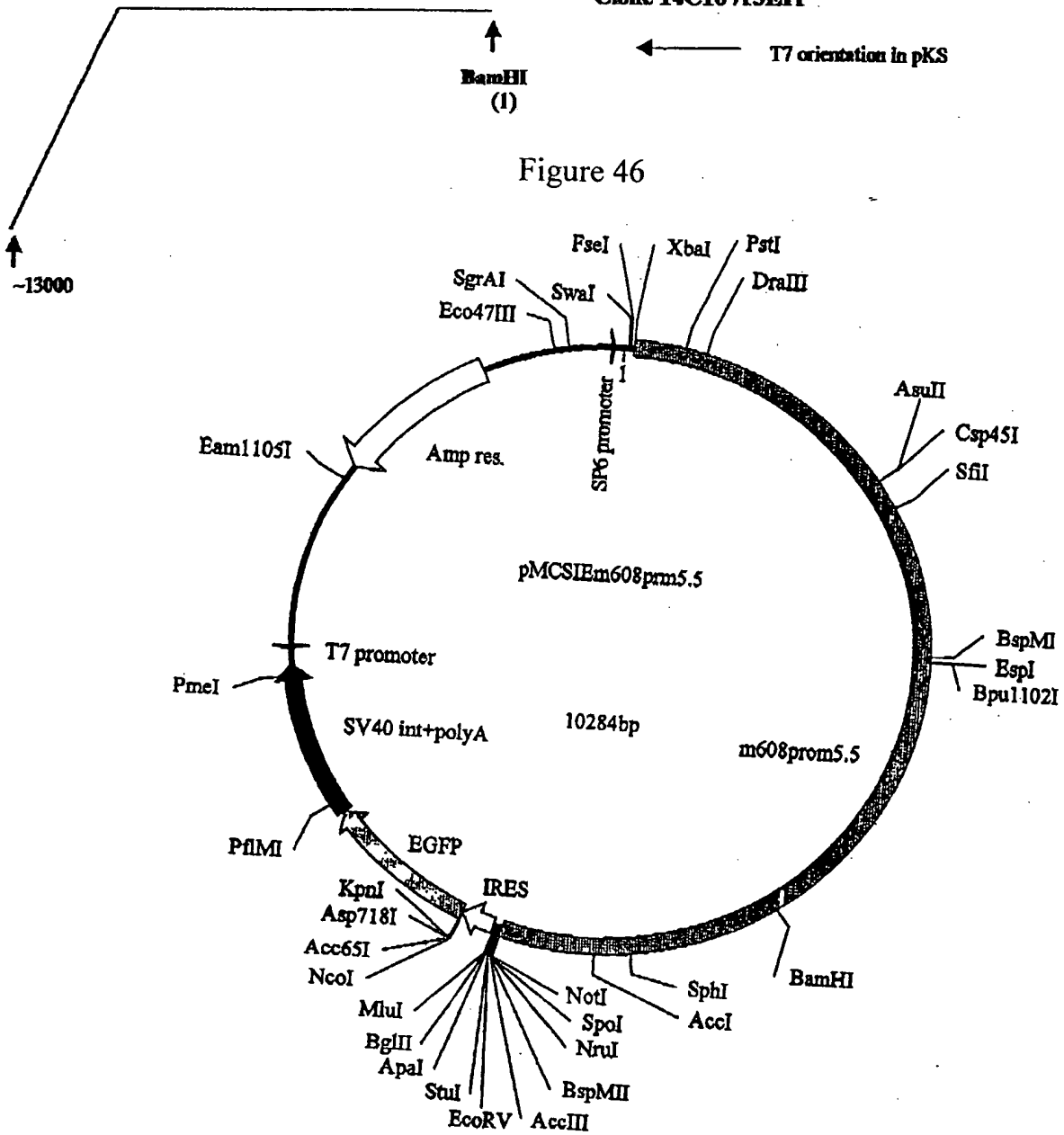
Figure 45



Clone 14C10 /15E11

T7 orientation in pKS

Figure 46



10284bp

Figure 47

↓ (XbaI)

TCTAGAAGTAAATGATCCTGAGTAGCGATCCTGGGAAAATACGTA
CTCTAACACACTGC
AATCATCTCTCTGTGGTTTGTCTGGAGCTGAGGTCTGGAAGGCTCGACCTTGGTTAGAAATA
ACCTACCGAATACAGAGCTATGACGTTAGTCTGGAAGGAGCTTTGGAAGAATGACAAAGCT
GTAGCTGCCAGAACATACTAGATGCCATATTTCCAAGGCAAGTGCCACATGCGGACAT
CTTAAGAATATGGTTGTCTCTGCAGTGCTAAGGACCTTGTTCGTGCCACACAGGTCTCCAG
GGTTAGTGCTAACTCTGACTGCTTGA
CTTTAATTCTACCTTGATCATTAAATGACTAGAA
ATCACTTGGTGATTAGCAACTGGATATGGAATATTACTAATTTGTACCCAAGCCAGGCCAC
CTCAGCTTTGGCAGCTCCATTCTGTGGAGCCCAGTCACGTGGGTTTGAATCAACTGT
ACTGTTTCTACTTACAAGACGCATTACCTGAGATGAGTCATTTTCTCACAAGCTTTTTTA
GAAGAGTCAATTAGACATATTCTGATGAAGTAAGCATATAAAGTGAGAGCAGCATGAATG
TGTTCCATGTATGCTCATGGATGCTATTATAATGTGGAAATAAACTGACTTTAAAAA
AGCTTATGATACTTGTACAGAGTAAATCTTCCATAAATATCATCTGCATTTATAAATTAT
TTTCATAATCCATCAATTA
AAAAACCTTAGAAATTTTGTTAACACAAAGATCCCTAGGCC
CTGCCCTAGGATGGTCTGTATGGTGGGCCCTGAGAGATGGAGCTTAAGA
ACTTACTTGCTCC
AGGAGCACATCTTCAGAACATCTGCCTCAAAACATTTATCCCAAATGCTCATCA
AAAGGCTC
ACTCACATGTGCTTCAACACAGGGATTAAACAGTCATTTTAGTCACATTTCTCAA
ACGGT
GGAAAGCCTGCTAGAGGAACAGGATGTATCAGGATAACATCCAACCTTACAAAAGGATGTC
ATAACCCCTACCCACAACAAACAAACGACAACAAACCCATAAAAAATTATCACGGCAAAAT
GAACTAAGCCATATGCAGAAAAAGTATTATATGTTCTCATTTGTGGGGTGT
TTTTCTTAAT
AGTCAAAATATGCAGAAATATAGACAAAGATGGTTATGCAAGTGGGGATGGCGAAGGATA
CTTGATAGATTAGAGGACACAAAGCAACACTACAGAGTGAAGTAATCCAGAGACTTAATG
TATAATATGAGGACTGTATTTAATAATTCTATTTAAGATACACAGCAAACGAGTGTATCTT
ACTAACACACACACTTACATAGAGAGAATAAAGTGATAGATACGTTTGT
TTTATCTTCATG
TAGCTGATAATTTCATATTGTACACCTCAACATAGATAACCAACAAAGAGGAAGAGGAT
AGGTGCCCTCTCCAGGGCGGAAGAGTACATTCGAAAGTCAGACACCATTGTGTAGATGTA
CCACATGGAGGAGCTAGAGAAAGTAGCCAAGGAGCTAAAGGGATCTGCAACCCATAGG
TGGAACAACATTATGAGCTAACCAAGTACCCCGAGCTCTTGACTCTAGCTGCATATATATC
AAAAGATGGCCTAATCGGCCATCACTGGAAGAGAGGCCCATTTGGACTTGCAAACTTTAT
ATGCCCCAGTACAGGGGAATACCAAGGGCCAAAAAGGGGGAGTGGGTGGGCAGGGGAGTG
GGGGTGGGTGGATATGGGGGACTTTTGGTATAGCATTGGAATGTAAATGAGTTAAATAC
CTAATAAAAAATGGAAAAA
AAAAAAAAAAAAAAAAAGGAAGGTCAGACACCTCAC
TTCACTGCTATCTCAACTTGCAAAACAGAAGGGGAGTCACAAACCCAGGACAAACCACAGT
GATTGAAGCGTCTTTGAATGTTATTGCTGTTGTTGTTACCACCATCATTAGCATATATTCAT
TGTGAAAACTTACGGGGTCTATGACATGTTTTTTTATTCAAGTATATCACATGCTGTCAGC
ATATTTGGCACCACTACAGCCCCAGCCCCCTTTGCCCGCCCCCAACACACACACACACA
CACACACACACACACACACACACACACACACACACACACCTTTACCTTCTCCTGGG
CATCATCTGCTCACTACCCACCCAAGCTTAATCCTTTTCTTCCCTGCAATAGTACCTCTC
CTATTTTATGTCTAGGTTCCCTCCCTCTGTTAGGAGATGGGAGAGGTCACGAAAGGAA
AGAATTTGTAGCCCTGAGCCAGCCCGGCCACAGAGCCTGCCACCAGACAGGAAAAAGCC
CAGGGCTTACCAGCACAGGAGGAGCAAACTCGCAGGCGAGCCTGGGTTGGCGCTGGTGGT
CCCGGGTCGATGGCCCGCCATTTCCAGAAGCCGAGGCTATAGCTGCGTCACCTGCCCG
COCTCCTCCCGAGTGAAGACCCCTAGAGGCTGAGCAGACCCCAAGGCGGTGCAATTCCA
TTGGCCCAAGGCAGAGGTGAGCGGCTGCTAATCCCTCGGGAAGTGAAGGGACCCAGAG

Exon1

AGTCTGGTAGATGTGGGAGCTGGGGTTACAGGGCGAGACAGAGGGTGGGATGGGCAGAAG
GGTCCAGGAAAAAGGAAAGTACTGGAGGGGAGTTGGGACAAAAGCAGCGACCAAGGGAA
CATCGCTTCAGTGACTGAAGCCAGGCAAAAGGAGCGGGAAGGATTATATGTAGCCTGGGA
CGCTTTCATAAACACTGATGACGTGTTTGTGCAAGCAAGCAATTGAGGAGAAACGCCT
GGGACGTGGGAAAGAAGGAGTGATCGATTAGTACTTGTAAGTTTAGGTGAGTTTGAGAAC
TAACCTATACTATTGAGGGAGAAGGAAGAGCATTCCAGCAGCAGCAGCAGCAGCAGC
GCAATCAGATAAAGGAAGCTTTGGTTAGTTTGGAAATGTATGATAACCATTAATAACA
GAAGCGCCTCCAGTTCTCTGAAGAGTCAGTCCCCAGCTAGTGAAGACTAAGCCTACTAA
GCCTTTTGTCTCCCTTGGGAAGCAAAAGCGTTCTTCAATCAGGTGAAGGCTCTCCTCAGA

AGATTTCTCTCTCTGCTTATGTTACAAGAGGATTCAAAAAGCAAGACAGAAGAGCTCAGG
 TATTGCCAACTCTTTTGTAAATACAGTTTGAGGCTTAAGTGACGGGAACATCATGTGGTA
 TTCATTTACGGCTCTCTTCTCTTATACTAACTCTTAAGGTGCATATAGTCTCTTCTGTTTCC
 CAGCTACCTTGTACCATCTTTGTTTATCTAATAATAGCAAGCTCATCTGCTTTTAAATCATC
 ACGCAGAGAGTATTCAAAAATATTCAGTGAATGTAACAGTGACAGTGTAGGCATAGAAAGTA
 ATCATTAGTAAATCTTAATTTGGGTAAACTCATTATACAGCTCCAGGTTGGGAGGGAT
 CACTGAGCCTTCGCCACGTGCGGGTTAAAGATATTTTCTAACAAGAGAAGCAGAATTCTTC
 CTGGCCATGCTCCCCATCACTGTGTCAAGTAAAGCAGAGGGGTGTTTCCAAGCAGAGAAAG
 AGCAGACAGTGTATGCCTGCAAAGTCAGAGACTCAGCCCTCCAGCTGGTCAGTTTACT
 GTCCTCCCGGTCATTAGTTGGCTCTGAAAAGGCCCATGTGTCCTTATTGGCAAGGACTTGC
 AGACATGCTAGAAAGAAATTTGACCTTTTTTCTAGTGGGTATTACAGCTGTAAAAGTAT
 TTTGGAAGGTTAAGCCAAATAAATAAACACATATTAATAATACAATGTTACAAAAAT
 GATCATATAAAGAAGTACATTTCATAAATGCAATGTGAAAAATATATATAATTTTATCTAT
 TTAAGTGTGCAAAGTTTCTAAATTCACATGTACTATTTTATATTTATAAAAAATATTTT
 AAAATGTATATAAAGTGTAAGGCTCTTGGTCAAAACAAGAGAGTTAAATTTACAACT

↓ (BamHI)

TTAATTGTCCCGATAACATTATTATGATCTCTAATGACAGGGATCTTCTTTTTCATTGGGA
 AATGAGAAGCTATGAAGATATGTTTACAATAAAGCCCATTTAGTGATAAAGTCCAATG
 GGAAGCTAGCACACACTGGTTTATAAAGAGAACAGTTTCTGAGTCTATGCAAGTTTACA
 CTCTAGGGAATAAGAGTTCCTCTTTCTCCAGATTTCAGTATGATTTGTTGTCATCTTTATC
 TTCTTGATGATGAGCATTATAAGTGAATAAGATAGGATCTCAAAGGAATGTCAATTTGG
 ATGCCCTGAACAATCTTTCAGGTCTTTCTTTCAGTTCACTAGTCTATTCATTTATTGGATAA
 TTGGGGGATGGTGTTAATTTTTCAGTTCTTATGGAATTCCAAAAACAAAAACAAAC
 AAACAAACAAAAACCTCTGAACTAGAACTACCAATCCATTACTGGGTATGTAACAAAG
 AGAAATCTGCACAGAATTTATTGCTACATTGTTCAATTATTCACGACAGCCAAGAATGTGGA
 ACCAACTTACGTAGCCGTCAAAATATGAACGGATAAAGAAAAATGTGGAAATGTGTACAAC
 AGAGTCCCATGTGGCCATAAAAGAGTGAAATCATGACATATGCAGGAAATGGATGCAACT
 GGAATCAATTGGGCTAATCAAAACAAGACAGACTCAAAAAGGAAACACCGTGTAGCTT
 CTCTGACAAACAGAAGCTAGATTACACTTGTACGTGCGCATGTGTGTTAGAATTTTATT
 TAGTTATACACTATTCTAATCTGTGAGTGTGTATAAAGGCATGCATGTAAAGCAAAAACA
 AGCTAGCTGGGGTGGGTAGGAGAGAAAGCAATGAGAGGAGTTAATAAGAACGAAGCATA
 GTAACATAGGTGCCAGGATGAAATGCATTAATTTGTATGCTAACTAAACCACAGACAGGA
 GGCACACGTTCAAACCAGGGTGAAATCCCAGCACAGAGAAGGGGAAGTAGACACAAAGT
 TTCGCCACTAACCAAGAAGCCATTTGCAAGTTGCTGCCTGCTGGGAGGGGGCGTTCCAGTTT
 CTCCAGTCTGACACTGTGTATAACAACAGTTGACAATACAAAGTTGGCATGATGGATGG
 TTTTGTGCTATTTTTCATTTTCTTACTGTTTGTGTTGTGTTGTTGTTGTTGTTGTTGTTG
 GCTGTGTTTTCATTTGTTTCTTTGAGAGAGAGAAGGAACATGAAATTTGGTGGGTAGGA
 AGCTGGAACGATCTGGAAGAAGTTGGGGAAAGAGAAAAATTGTATGGAGCATATTTAA
 ACAAACAAACAAACAAACAAAGGTTCAATTTGCCACAAAAAGGTGTGAATTAATTAAC
 CAGTTACGACTCTTAAAGAAATATTTCCCAATTATTTCCAGAGTTGCTATGTATGCTGTGC
 CTAGGACTTTGCTTGAAGTGGCCCTATAACTCTGGTGTGGTGTCTTTTCAGG - Artificial
 NotI site

(SEQ ID NO:17)

CACAGACCTTCTCTCTCTAACTCTCTCCCCCATCTTGTGTGCTTCATCCCAGACTTCAACAC
CAGCAAGCACACTCTGCTAATGCAAGGGCTGCTCCTGTCCAGGACAACAAGGAGGCTGAAG
GCAGACCCACACGTTTCCAACTGCTCCTGAGAGTCAATCCCCCTAGACTCATCTATAGCAG
GAAACCTGCTGTGATCTCCATTTCTTCTGTACCACATCCCCAAGTTATCACAAGGAGTTT
TCTTCAAACCTTTCTCTCCAGCAAAACCCCTTCAGCTCCTTGGGTACTTTCTCTAGCCCT
TCATTGGGAACCTGTGTCTCCATCCAAATGGATGGGTGTGAGCATCCCACTTCTGTATAGAAT
CTTGGTCAGTGCAGTCTTTTGTATCCTCAAGAACACTGGGTCTGAAAATTTTAACCCAAAG
AACTGTTTTTTGTTATGATTGCTGCAATCTCTTTCAATTCCAATAAAGAGTAAGCATCTCAT
TCCTTTGTCTCCTCCTTTTCAGTACCACCTGCGCTTGTCTGCCTTTCTCAAAGAAATCAATAAA
ACCAAAGTGATATAGATTTCATGGCATTCTCTAACTGCTACATCCACTCCAGTAGTATCTC
ACTTGGCAGGTGTAAAGGCTGGAAGCAGTCAAGGAGCAGTTTCACAGAACTTAGCCCTC
CTGGAACCTTGGCATTCCCATAGCTAGAATTGCCCGAGATTGTGCTGTAGATATTGTGGTG
GGTCTTGCATGCTTTCTTGCAGTATTTTACTGGATAAGAGTTAGAAAATCTCAGGGCGAGCT
TAGCAAAAGTATACCTAGAATCTTCATGACAGTCAGGTATTGCAAACTACATTGCATATTA
GAAGAAAGTTGGTAAATTTCTTCTGACAAATGGAGATTCCCTACAGATAACTTAAAGAAC
AGCTAAGTCACACTCATATGCAAGAAATTTACCAAGGCCTAGGAAAGGGGGGGGGGTACT
GCTTTATTCTAGTAAAGGCTGCTAGAGCAGAAACCCCTGGGTCTAGCTTTTACAAGGTTT
AAGGTTGTAGCATAAAATTTGTGACTAGAGTGTGAAATCTTTACCTGTATTAGCTGACTCTA
GGCAGAGCTGTTTTATCTTTACTGTAAACATTACCTGGTTCCTGTGAGTCTTTGAAGGCAT
TCTCTGTTTTGTGACAGATACTTCTATGTACCTCGCTGCTGTGACACCCCTACTCCTTTGT
TTCTGTATTATATAAGCCTGGTGTTCCTTTGTGAAAAATTACATCCAGATACAGCACTC
CCTTGTGTCTGTGTCTTTTGTCAATTTCTGGCCAACTCCATGCCACCTGCCAGAACCCCTA
GTCTTTTCCACAGATTGAGGGAGGCGGAGCTGAGCCTGGTCCATGGCATCTAACCACTGTCA
GCTCACTGTTGGTGACTACCTCAAGGTACAAGCTCCATTACTAATGAAACAAAATTAGAT
AAGTGTGGGTCCAGGAAGCAGGTTGTACACCTGTCTGAATGAACATTATGAAATGACTG
AAATAAGTTAAACCATCTCTTCTCGTTTGCTAATATAGCAAATAAACCGAGTTTCTGAGC
TGCTGCTGGTGTGTCTCCATCAGAGGGCAGAGCCAGTCTGATCCTAGCTTTCTGTATGTG
TGTCATTGTTTCTTCAGTTCCTGTTGCCCAATTAGGAAATCCTAAGCCATGAAAGCCATG
AATCTGGGAATGACTTTTCTAAGAAATGCCACGTGAACCTTGCCTTTCAACGTTTTGCCTG
TAAACAGATATATGTTGCGCAGTTTATAATCAATAAAGCTTTGAAATAATATAAATCACTC
CATTTCTATTCTGCTTCCACGCTGAGCATCTGTGTTTCCCAAGGGACCAAGAGCATTTGA
AAAGTAGTGATTATGACCTGCTTTGTCTGTACTATAAAAGCTTCATGAAAGGGCAGCC
ATGTTGAAACATGGAACCTTGGGGTGACCTGTATCTGTGTTCTGGGTGCTGCTCACTCATA
TTTGTCTCCAGAATAAATGAGTTTATCAACTTCGAGGAAAAAGTTGTGTGTTTGTATAGCA
CGCCCGTGAGTCCCAACCTTACTTCTCTGTAATCTGTATATGGTAGAAAAAGTTAATTT
ATGTGATTTCTTCCAACTCCAAATATTTCAAATCTTTTAGCCCTCAGCCTGGGATTTCTTG
ACTAAGTCTATTGATTGGGAAGATCTCAGTGGTTAGGATTTGCAGTCAATGATGTTACATG
TCAGGCTAAAGCTGAAAAATATGACAAATGAAATGTCAAATGTCAATGTGCTGCGCTGGGAATGTG
AGTGTTAGGGGGTTTTAAAGAAACAAATACCTACTCTAAATAGTTAATAAGTCCCATGGTT
CTATTCTAGTTTTGAAATAATGTTCCCTAGTATACAGCAATTAATTTGAAATGAATAGCTTC
TTATCTTGACCAATCTCAGTGACTTCATCCGTCCCAAGTCAATGTTTTCATATTTCATAAGGAT
AGGTCTCATACACCACTGTTTATCATTTTGGGATCTGCATTTTTCTGATGCAAAATGATTT
ATTTCTCAGAGCACTGGAATTTGGGTTGAATCATCTTATAACGGGCCAAAACTAAATGCTTC
TGTGCTAAACAGAGAGTTACAAGACCTTTTTATGTGGATGGCAGCAATTTAGTCAATCCTTA
TGACAGAATGTGAGAGTGGAGCTCCCACTGGGGGAGGGGCTGGTCTTGGCAGGATTCTC
TTGGAACATCACACAAAGAAATTCCAAATTATGAAATGCACATGATCCATCCAGAATGTG
ACTTTTGAACCTTTGAACATGAGCTTTTAAAGTACGTTTGGGTGTTTCAGACCTTGACTTTGAG
GTGAAGGAAGAGCTCGCCAACTCCTTTTTATGTAAACAAATATATCAAGATCTAATGTGA
GACAGTATGCCAGTCCCAAGATCTGTCAATGATGACTGAAGACACATTTGCGATGTTATCACT
AAGGCAGGAGAAGGCAAGCTACAGTGAAGCCAGTTCACATAAAGCTTTTATGAGAAAT
AGATAAGAAGGGTTTCTAATTTTTAAATTTTTTTTATTAGATATTTCTTCATTACATTTT
AAATGTATATCCCAAAAGTCCCTATACCTCCCCCTCACCTGTCTCCCTACCCACCCACTC
CCATTTCTTGGCCCTGGTTACTTGTGATAGTGGTCAATGATGATCCCAAGCTTTTACATGTCTC
ACTATCTGGTCTATTGCAAGAATGGCTGCGGAGCTGATGCGAGTCAAGATACAGACACCTAC
AGCCAAACAGTGGGAAGGAACCTTGGGGACTCTTATGGAAGAAAAGGAGGAAGGGTTATGG

CCCCGGATGGGGAAAGGAACTCCACAGGAAGACCAACATACTTGGTCAACTAACCTGG
ACCCTTGGGGCTCTCAGAGTCTGAACCAACCAACCATAGAACATTTCATGGGCTGTACCCAG
GCCTCTCCACTCATATGTAAACAGATATGTGGCTTGGCCTTCATCTGGGTCTGAACAACTA
GATGGGGGTTAGGGGTGGGGATGGGGGTTATCTCAAAAGCTGTTGCCTGTATGTGGGATA
TGTTCTTACTGAGCTGCCTAGTCTGGCCTCAGTGGGAGAGGAAGCACCTAGCTTTAGCCTT
GTAAAGACTTGAAGTCTGAGGTGTGGGTGGAGGGTATACTCAGGGAGGCCCTCAOCTGC
TAGGAAGAGAAGAGGAGGGGGAAAGACTTGGGGGAGGGGGCAGTGAGCAGGTTGGTAAA
ATGAATAAGAAAAAAAAAATAAAAAATAATTAAAAAAAAAAAAAAGAATGGCTGCTGAGCCC
TACTCTAAAAACCATTGCATCCCCCCCCCCCCAATCATTTCAGTGACTACGAATTAAAAATCATT
GATACTAACAATAGATGTAGGAAACTATTGTTAACTTCTTTGTGAECAGAGTGGTATTG
GAACCTTTTTTATTGAAGCTTTCACACAGAGCCTTGTTCTTTTCATTTCCCTGTACATGCATG
TAGCTTAATGATGTTTCAGTGAATTAATAAATAACAATGAAGAATAAAGACAACCTGTATTTT
AAGGATTCTTCGTATATATTTAAAAATCTAAGGTGGTCACTGGAAGAAATGTCTTCAGTT
TTTCTATATATGTTTACTCTATCGTATGTTAATTAAATTATATGCAATAATTCATAAAATCTA
CAACATAGTATGTAACCTATAAGAAAAGTAAAAATTCATGAAATTGTGAAGGTTACTTTTC
CTTACCCTCAGAAAACACTGGGTTTGAATAATTCTTATTTTGGTATCAGTGAAGAATTGAA
AGAATGTAATAACCTACTAAGGCAAAACATAGAAGTTGAAATTAAAAAGAGTAGACAGGA
GAAGTAATAAGGCAAATAATGAATATTGCTTTAAATAGTTCTTAATGTATCATCTAACTA
GGGTGTGATTCTCCAGACTTGACTCCATCCAAAATATCCAAAATGACTCTAACCACAGTCA
TTGAAACAATGTGTTGAAAATAATAAACATTTCOTACTTGAAAATTTCAGATTTCTCCTACT
TTGCTTTTTATTGCTGTGATAAGCACCATGACCAAAGCAGCTTAT

(SEQ ID NO:18)

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

72/85

Figure 49

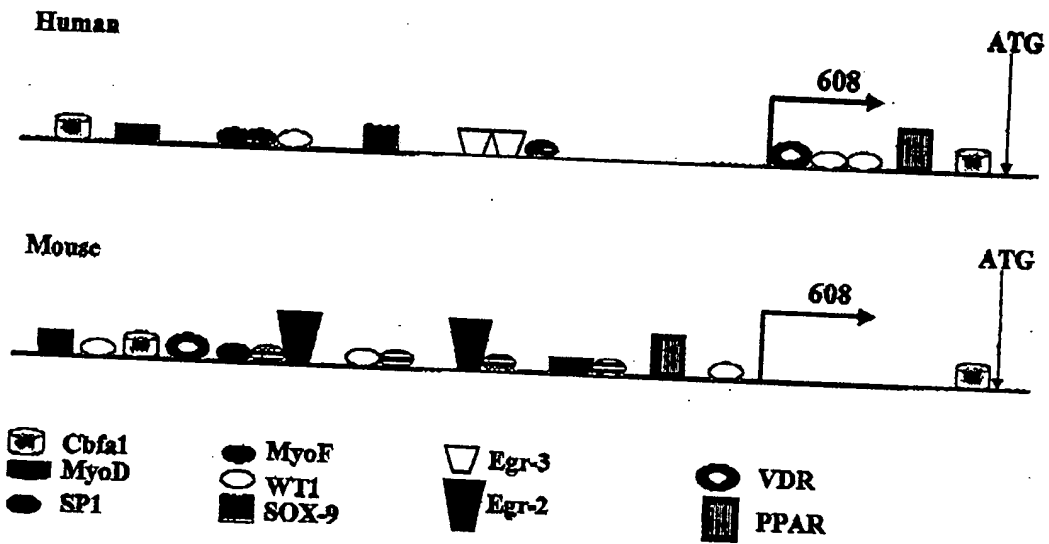


Figure 50

TAAGCCTTTTGCTCCCGTTGGAAGCAAAGAACGTTTCCTTCAATCAGGTGAAGGCTCT
CCTCAGAAGATTTTCATGTCTCAGCTTATGTTACAAGAGGATTCAAAAGCAAGACAGA
AGAGCTCAGGTATAGCCAACTCTTTTGTTAAATACAGTATGAGGCTTAAGTGTACGG
CAACTCATGTGGTATTCATTTGCGGCTCTCTTCTCTTATAACTAACTCTTAAGGTGCA
TATAGTCTCTTCTGTTTCCCAGCTACCTTGCACCATCTTTGTTTATCTAATAATAGCA
AGCTCATCTGCTTTTTAATCATCACGCAGAGAGTATTCAAAAATATTTCAGTGATGTA
ACAGTGACAGTGTAGGCATAGAAGTAATCATTAGTAAATCTTAATATGGGTAAACT
CATTCATAACAGCTCCAGGTTGG (SEQ ID NO:19)

TTTGGAACCAACCCAGATGCCCTCAACAGAGAAATGGGCCAGAAAATGTGGTCCA
TTTATCCAATGGAATACTACTCAACTTATTA AAAACAACGACTTTCATAAAATTTT
GGCAAATGNATGGTCTGNAGGATCTTGAGTGAGGTAACCCAATCACAAAAGAACAC
TCATGGTATGCACTCACTGATAAGTGGCTATTTGTCTATGGAGTGATTTAAAAGGGA
AGAAGACACATAGCTTTTTGTGTGTATAATATTAAGATGGAAATTTGCCAGTGCTGT
TTGGCTTATGAGTGAATCTTGTTTCAGTGGATTACCGGAAGAAAATAATAAGTGAAC
TGTAGGAAGAAGTAGTTAATCAAGGTGACAAAGTATCCTGACACATTGGGAAAAGA
CCACAGTCCAGGAACTGAGTCTTAAGGATTCATATTA ACTCCAGTTCCCCATGTGC
AGCTCTGAGACTTTGGCAGATCAGACACTTA ACTTCACCAGCTTCCTACACAGAGCA
GTTACTATCCTTGCCCTTCACACATGGAGTGTGCCATTAAGTGCCTGAACATGAGTCT
GACTTGTTAATAATCTTTAAAATCCAATTGTGTGTAAAGTATGTGACCAAAGAGCAT
GGTCATGCTATTAACCTTTGATGTTCTATGGACTCTTAATTTTATGGTAGAAATGTCA
ACAAGCTTGTGGAGGCTGGAAGATACAAGGCTTAAGAGGATGGCCTTTCAGTTTTG
AAAGTAATTCAGTATGTGTTCTGGCATCCCTTTTCTAAAGCAATTTAACCCCCCAA
GTAGGCATAATTTTAATGCTTACTTCATCAGAATATATCTAATTGACTCTTCTAAAAA
GACTTTGGTATGCATAGGATCTAAATGTAAATGTGATTTACTGACATAATAAATAGG
AGAACTGAGCTAGAATAGGTATAAAATATGTGCTGGCTTTCTAATAGGTCTTATAG
GTTATATAAGAGGTGGGAAAGGAATATTTGAAACATCTAGAAGTAAAATGATCCTG
AGTAGCGATCCTGGGAAAATACGTA CTCTAACACACTGCAATCATCTCTCTGTGGTT
TGCTGGAGCTGAGGTCTGGAAGGCTCGACCTTGGTTAGAAATAACCTACCGAATAC
AGAGCTATGACGTTAGTCTGGAAGGAGCTTTGGAAGAATGACAAGCTGTAGCTGCC
CAGAACATACTAGATGCCATATTTCCAAGGCAAGTGTCCACATGCGGACATCTTAAG
AATATGGTTGTCTCTGCAGTGCTAAGGACCTTGTTTCGTGCCACACAGGTCTCCAGGG
TTAGTGCTAACTCTGACTGCTTGACTCTTTAATTCTCCCTTGATCATTAAATGACTAGA
AATCACTTGGTGATTAGCAACTGGATATGGAATATTACTTAATTTGTACCCAAGCCA
GGCCACCTCAGCTTTGGCAGCTCCATTCTGTGGAGCCCAGTCACGTGGGTTTG
AATCAACTGTACTGTTTCTACTTACAAGACGCATTACCTGAGATGAGTCATTTTTCTT
CACAAGTCTTTTTAGAAAGAGTCAATTAGACATATTCTGATGAAGTAAGCATATAAAG

TGAGAGCAGCATGAATGTGTTCCATGTATGCTCATGGATGCTATTATAATGTGGAAA
TAAACTGACTTTAAAAAAAAAAGCTTATGATACTTGTACAGAGTAAATCTTCCATA
AATATCATCTGCATTTATAAATTATTTTCATAATCCATCAATTA AAAAACCTTTAGAAA
TTTTGTTAACACAAAGATCCCTAGGCCCTGCCCTAGGATGGTCTGTATGGTGGGCC
TGAGAGATGGAGCTTAAGA ACTTACTTGCTCCAGGAGCACATCTTCAGAACATCTGC
CTCAAAACATTTATCCCAAATGCTCATCAAAGGCTCACTCACATGTGCTTCAACCAC
AGGGATTAAACAGTCATTTTAGTCACATTTCTCAAACGGTGGAAGCCTGCTAGAGGA
ACAGGATGTATCAGGATAACATCCAACCTTACAAAAGGATGTCATAACCCTCACCA
CAACAAACAACAACGACAACAAACCCATAAAAATTATCACGGCAAATGAACTAAGC
CATATGCAGAAAAAGTATTATATGTTCTCATTGTGGGGTGTTTTTCTTAAATAGTCAA
ATATGCAGAATATAGACAAAGATGGTTTATGCAAGTGGGGATGGCGAAGGATACTT
GTAGATTAGAGGACACAAAGCAACA ACTACAGAGTGAAGTAATCCAGAGACTTAAT
GTATAATATGAGGACTGTATTTAATAATTCTATTTAAGATACACAGCAAACGAGTGT
ATCTTACTAACACACACACTTACATAGAGAGAATAAAGTGATAGATACGTTTGT
ATCTTCATGTAGCTGATAATTTTCATATTGTACACCTCAAACATAGATAACCAACAAA
GAGGAAGAGGATAGGTGCCTCTCCAGGGCGGAAGAGTACATTTCGAAAGTCAGACA
CCATTGTGTAGATGTACCACATGGAGGAGCTAGAGAAAGTAGCCAAGGAGCTAAAG
GGATCTGCAACCCTATAGGTGGAACAACATTATGAGCTAACCAGTACCCCGGAGCT
CTTGACTCTAGCTGCATATATATCAAAAGATGGCCTAATCGGCCATCACTGGAAAGA
GAGGCCATTGGACTTGCAA ACTTTATATGCCCCAGTACAGGGGAATACCAGGGCCA
AAAAGGGGGAGTGGGTGGGCAGGGGAGTGGGGGTGGGTGGATATGGGGGACTTTT
GGTATAGCATTGGAAATGTAAATGAGTTAAATACCTAATAAAAAATGGAAAAAAA
AAAAAAAAAAAAAAAAAAGGAAGGTCAGACACCTCACTTCACTGCTATCTCAACTTG
CAAACAGAAGGGGAGTCACAAACCCAGGACAAACCACAGTGATTGAAGCGTCTTTG
AATGTTATTGCTGTTGTTGTTACCACCATCATTAGCATATATTCATTGTGAAA ACTTA
CGGGGTCTATGACATGTTTTTTTATTCAAGTATATCACATGCTGTCAGCATATTTGGC
ACCACTACCAGCCCCAGCCCCCTTTGCCCCGCCCCAACACACACACACACACAC
ACACACACACACACACACACACACACACACACACCTTTACCTTCTCCTGGGCA
TCATCTGCTCACTCACCCACCCAAGCTTAATCCTTTTCTTCCCTGCAATAGTACCTC
TCCTATTTTTTATGTCTAGGTTCCCCCTCCCCCTGTTAGGAGATGGGAGAGGTCACGA
AAGAAAGGAATTTGTAGCCCTTGAGCCAGCCCGGGCCACAGAGCCTGCCACCAGAC
AGGAAAAGCCCAGGGCTTACCAGCACAGGAGGAGCAA ACTCGCAGGCGAGCCTGG
GTTGGCGCTGGTGGTCCCCGGGTCGATGGCCCCGCCATTCCCAGAAGCCGAGGCTATA
GCTGCGTCACCTGCCCCGCCCTCCTCCCGAGTGAAGACCCCTAGAGGCTGAGCAGAC
CCCAAAGGCGGTGCAATTCCATTGGCCCAAGGCAGAGGTGAGCGGCTGCTAATCCC
CTCGGGAAGTGAAGGGACCCAGAGAGTCTGGTAGATGTGGGAGCTGGGGTT CAGGG
CGAGACAGAGGGTGGGATGGGCAGAAGGGTCCAGGAAAGGAAAGTACTGGAGGGG
AGTTGGGACAAAAGCAGCGACCAAGGGAACATCGCTTCAGTGACTGAAGCCAGGCA
AAAGGAGCGGGAAGGATTATATGTAGCCTGGGACGCTTTCATAAACACTGATGACG
TGTTTGTGCAAAGCAAGCAATTTGAGGAGAAACGCCTGGGACGTCGGAAAGAAGGA

GTGATCGATTAGTACTTGTAAGTTTAGGTGAGTTTGAGAACTAACTAACCTATACTA
TTGAGGGAGAAGGAAGAGCATTCCAGCAGCAGCAGCAGCAGCAATCAGATAA
AGGAAAGCTTTGGTTAGTTTGGAAATGTATGATACCATTAAAATAACAGAAGCGCCT
CCAGTTCTCTGAAGAGTCAGTCCCCCAGCTAGTGAAGACTAAGCCTACTAAGCCTTT
TGCTCCCGTTGGAAGCAAAGAACGTTCCCTCAATCAGGTGAAGGCTCTCCTCAGAAG
ATTCCTGTCTCTGCTTATGTTACAAGAGGATTCAAAAAGCAAGACAGAAGAGCTCAG
GTATTGCCAACTCTTTTGTAAATACAGTTTGAGGCTTAAGTGTACGGGAACTCATG
TGGTATTCATTTACGGCTCTCTTCTCTTATACTAACTCTTAAGGTGCATATAGTCTC
TTCTGTTTCCCAGCTACCTTGTACCATCTTTGTTTATCTAATAATAGCAAGCTCATCT
GCTTTTTAATCATCACGCAGAGAGTATTCAAAAATATTCAGTGATGTAACAGTGACA
GTGTAGGCATAGAAGTAATCATTAGTAAATCTTAATTTGGGTAAACTCATTCATAA
CAGCTCCAGGTTGGGAGGGATCACTGAGCCTTCGCCACGTGCGGGTTAAAGATATTT
TCTAACAAGAGAAGCAGAATTCTTCCTTGCCATGCTCCCCATCACTGTGTCAGTAA
GCAGAGGGGTGTTTCCAAGCAGAGAAAGAGCAGACAGTGTTATGCCTGCAAAGTCA
GAGACTCAGCCCTCCCAGCTGGTCAGTTTACTGTCCTCCCGGTCATTAGTTGGCTCTG
AAAAGGCCCATGTGTCCTTATTGGCAAGGACTTGCAGACATGCTAGAAAGAAATTT
GACCTTTTTTTCTAGTGGGTATTACAGCTGTAAAAGTATTTTGGAAAGGTAAAGCCA
AATAAATAAAACACATATTAATAAATACAATGTTACAAAAATTGATCATATAAAGA
AGTACATTCATAAATGCAATGTGAAAAATATATATAATTTTTATCTATTTACTGGTGC
AAAGTTTTCTAAATTGCACATGTACTATTTTTATATTTATAAAAAATATTTTTAAAATG
TATATAAAAGTGTAAGGCTCTTGGTCAAACAAGAGAGTTAAATTTACAACTTTA
ATTGTCCCGATAACATTATTATGATCTCTAATGACAGGGATCCTGCTTTTCATTGGGA
AATGAGAAGCTATGAAGATATGTTTACAATAATAAGCCCATTTAGTGATAAAGTCCA
ATGGGAAGCTAGCACACACTGGTTTATAAAGAGAACAGTTTCCTGAGTCTATGCAA
GTTTACACTCTAGGGAATAAGAGTTCCTCTTTCTCCAGATTTCACTAGCATTGTGT
CATCATTTATCTTCTTGATGATGAGCATTATAAGTGGAATAAGATAGGATCTCAAAG
GAATGTCAATTTGGATGCCCTGAACAATCTTTCAGGTCTTTCTTTCAGTTCCTAGTC
TATTCATTTATTGGATAATTGGGGGGATGGTGGTAATTTTTTTGCAGTTCTTATGGAA
TTCCAAAAAACAAAAACAAACCAACCAACCAAAAAACCTCTGAACTAGAACTACC
AATCCATTACTGGGTATGTAACAAAGAGAAATCTGCACAGAATTTATTGCTACATTG
TTCATTATTCACGACAGCCAAGAATGTGGAACCAACTTACGTAGCCGTCAAAATATG
AACGGATAAAGAAAATGTGGAAATGTGTACAACAGAGTCCCATGTGGCCATAAAAG
AGTGAAATCATGACATATGCAGGAAATGGATGCAACTGGAAATCAATTGGGCTAAT
CAAAACAAGACAGACTCAAAAAGGAAACACCGTGTAGCTTCTCTGACAAACAGAAG
CTAGATTTACACTTGTACGTGCGCATGTGTGTTTAGAATTTTATTTAGTTATACACTA
TTCTAATCTGTGAGTGTGTATAAAGGCATGCATGTAAAGCAAAAACAAGCTAGCTG
GGGTGGGTAGGAGAGAAAGCAATGAGAGGAGTTAATAAGAACGAAGCATAGTAAC
ATAGGTGCCAGGATGAAATGCATTAATTTGTATGCTAACTAAACCACAGACAGGAG
GCACACGTTCAAACCAGGGTGAAATCCCAGCACAGAGAAGGGGAAGTAGACACAA
AGTTTCGCCACTAACCAAGAAGCCATTTGCAGTTGCTGCCTGCTGGGAAGGGGCGTT

CCAGTTTTCTCCAGTCTGACACTGTGTATAACAACCAGTTGACAATACAAAGTTGGC
ATGATGGATGGTTTTTGTGCTATTTTTCATTTTTTTCTTACTGTTTTGTTGTTGTGGT
GGTTGTTGTGGTGGTGGCTGTGGTTTTTCATTTGTTTCTTTTGAGAGAGAGAAGGAAC
ATGAAATTGGGTGGGTAGGAAGCTGGAAACGATCTGGAAGAAGTTGGGGAAAGAG
AAAAATTGTATGGAGCATATTTAAACAAACAAACAAACAAACAAAGGTTTCATTTT
GCCACAAAAAGGTGTGAATTAAATTAACCAGTTACGACTCTTAAAGAAAATATTCCC
AATTATTCCCAGAGTTGCTATGTATGCTGTGCCTAGGACTTTGCTTGAAGTGGCCCTA
TAACTCTGGTGTGGTGTCTTTTCAGGATGCAGAAGAGAGGCAGGGAAGTCAGCTGCT
TGCTGATCTCCCTCACTGCCATCTGCCTGGTGGTCACCCCTGGGAGCAGGGTCTGTC
CTCGCCGATGTGCCTGCTATGTGCCACAGAGGTGCACTGTACATTTTCGGGACCTGA
CCTCCATCCCAGACGGGCATCCCAGCCAATGTGGAACGAGTCAATTTAGGGTGTGTG
GACCTTGCTGATCTCCTTCTCAGAGAGGGACCACTGATTTTCCTGGTACTTTGCCCC
CCAAACACCTGTGATTACTTTTAATAGTTTTCTTCTAAAATGGGTTTCATACAAACCTT
ATATTGTGGAGACAATGAACATTTTATCCCAATAGTCTTTTACTAGAAGTGAAGCC
CCTCTTAGTTGTTTGGGAGCCTCATAATTATGGGGCAGCTTTATTCTGAATGAATTTT
AAATGAAAAAGATACAGTTTCTGTTAACAATCATTATGATACCAAGGAAGAGGAAT
TGTCATTGAATATTTTAAAAAAGCATTTTCTTTTGCAATTTATAAATACCCATTACAAA
ATGGCTTACTTAAAATACTTGCCTTACTAAATCTGACAAATTATGGTGATATTTTGAA
GGTTTATGAAAATTTGTTTATGTGTATAAATGCACAAGAAATGGGATATGCCATCAC
CTATGTGCCATTAGTGAGCATGTACAGTATGCCAAACACTATTGTTACGTTTGGAG
GAAGTAATGGGGGTGGGGGAGCAACAAGGGTTATAACCGTATACCCAGTGCCTTGG
AAGCGATTGCAAACAGTAAAGACTGACATTGTGTTCTCCCTATGAGGGAGGGGCCTT
GGGCTGAGCACTTTGCAATGAGCATTGTGCTCATTGTGCTGGCAGGTTTTATGATAAC
TTGACCCAAGCTAGAGTCACTGGAGAGGAAGGAACCTTCAACTGAGAACATGCCTGA
AGAAGATCAGATTATAGGCAGGCCTGTGGGGCATTTTCTTAATTAGTGATTTCATGGG
GCAGGGCCCAGTCCATTGTTTCGTGGTACCATTCTCAGGCACTATTAACAAAAA
AACAGGCTGAGCAAGTGCAAGGAGCAAGTCAGTGAGCAGCAGCCCTAATGATCTC
TGCATCAGCTCCTGCCTCCAGGTTCTACCCTATTTGAGTTCCTGTCTAGCTCCCTA
CAGTGATGAACAATGATGTGGAAGTATAAGCCAAATAAATCCTTTCTTCCCCAACTT
GCTGTTGGTCATGATGTTTCATCACAGTGATAATAGTCCTCATGAAGATGCTGGTGT
TTATAACACCTTTGGACTAAATTCTGTTATCTATAGCTGAGGAAAATGGAGCATAGA
AAGTCTCCAGACTACACCAGAGTGTAATCTGGGCCTGAGCTTAGAATCACACCCAC
GTGCACTCCACTGCCGGGGCTTCTTAACCGGAACACAGTTGTAAAAGGGAATTTTCT
GTTTGTTCATTTTGACATGTGGACTTTAATTGACGATTCATCTGAAGCTGAAAATG
ATTTTTTTCCAGGTATAACAGCCTCACTAGATTGACAGAAAATGACTTTTCTGGCCT
GAGCAGACTGGAGTTACTCATGCTGCACAGCAATGGCATTACAGAGTCAGTGACA
AGACCTTCTCGGGCTTGCAGTCCTTGCAGGTGAGATAGGTAGAGGGTGATGGAGGC
TGAGAAGAGAGGTGCAACTGTGGGTATACCCAAAAGCTGCTGATTCCCGTGGGAG
ACATTCTATAAGCATTCTATAAACTAGAGGCAGATATCAAGGAAGGATTTCAATTGT
AATGCAATTTTATGAGAAAATTTGAATATTAAGAAAATGCTGGGGAAAATGCTTAC

ACAATTGCGAGGACCTAATTTAGGATCTCCAATAGCCACATAAAAAGCACAGCATG
 GCGGCAGACACCTGCAATTCCTGTCCCTGGAAGCACCTGTTTCAGAATCCCAGAGACT
 CATTGGCCAAACACTCTATTCAATCAATGAAGTCCATATTCAGTGACAAAACCTTGAC
 TCAGAAACTAATGTGGAAAGCATCAGGAAGACAGCCAACATCTGGTCTCTACTCAT
 GCATGAATAAGGGATCCCAGAGAGAAGGGAAGGAAAAAGGAAGGAAGGAAGGAAG
 GAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGAGAGGGAGGAAAGGAGGG
 AGGGAAGGAAGGAAGGAAGGGAAGGAAAAAGAGATGGGGAGGGAGGGAAGGAAG
 GAAAGGGGGAGAAAGAAGAGAAGAAAGGAAAAATAAATAAATTTTCAGGGATTATT
 ACACCTTTAAATTTTATCCATAAAAGGTCATTTCCACCTGTTTGTCTGGAAGTAGAGT
 GGGATCCCTTATATAAGGGCAGTCTTTAACATAGTAGCATTTTATAAACCATTACAA
 ATTTTGAGTTTTCTCTACTTTTTATCCTCTACCATCTTCAAACCTGAAACTACAATTATT
 CCCACAAATGAAGAAAATGCTGTAAGAGTTTTTCACACACCGAAGTGGGAAACTTAA
 GGATTAGACAAGTCTAACAATGAGAATGGGGAGAACAAAAAGAGACTGCACAGGG
 AGCCCTTTCTCTGCTTATAATCTTGACACTTGAGAAGCTAATTGACGCTGCATGACTA
 CTCAACTCTTTAAGCAAACAATGCTGTTGTTTCATGAAAAGCACATAAAGTACATAT
 GTCCCATAAATATTCATCAAAATTTGCATGCAGCACATAATAGCAATCAAAGCAATAA
 CACCCACTGTTTCACAGAGACTTTAAACATGAAACTGGAACATATGTCTAGTGTTTTGA
 CTTAGGGTACATAGTATGCTGTGTCTGTATGTACCAATGTTGATTTAGGTCATCAGA
 CAGCATTTGGAACATGTATCTTCAGGAGGAATCATTTCATGTATCCTGCATGAAATTC
 TCCACCTATGTTTATTCTCTTAGCCAGGTTTTTCTCTGATGGAGAAACATTGGGGTTTG
 AGGTTTTACTCCCAGGTAACATTTAGGGAAAAGCTGTCTATGTTCTCAGTTTGGCTTT
 TATTTATGAGGGATGTTGGTATTCCAGAAAATTCTCTTTTGAAGAGATTACAATTTA
 GGTCAAAACAGAAAAATATGTAAAAAGTTATTGTTTTTATTAGTATTTTCATGTTCTTT
 TCTTTTTTAAAAATGGTATGCTTAGAATAATTAAGATTAGATTAGATTAGATTAGA
 AAATAATCAGAGAGGGATTTGATGAATGCTAAAGCATCATGAAAAATTCAAAATTT
 TTTGCTTCTAATTCAGAATCAATTAATTCATATTACTATAAAAGACAGCACGCCAG
 ATGTGTGCCAGCTGAGGAGTGGATAAACTGTGTAACTGAGTGCTATGTAGAAACA
 GAAAGGAGTGAAGGGTTGATGTGCGCTGCAACATCTTGAAAACATTTCGGCTACATG
 ATGGAAGCCAGGCACAAAAAGCCACATATTGCATGGTTATGTTTATATGAAATGTTT
 AAAATACATGGATTCTTAGCAAACAGAGTAAGATGTTACTTAGGGTCAGGAAAAGA
 TTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA
 TCCTAAGACGAAATAGTTGAGGTAGATATAGTTATATCCCTGTGGATATTGTAATAA
 ACCAGCATGCTGTGCTCTGAGAAGGGCCTAATGAAGGGGCAGGAGGAAGTGAAATG
 AGATGGTAGAAAGGAAAGTCATATACCATGGCTTCTCTCGTGGGTGGAATCTAGAT
 ATGTTAATATATTGACATAAAGGAAGGAATTGTTTAGGGAAGGATCAAAACCAACA
 GGAGTGAGGGAGACAATAGGAACCAATGAGAGGCAAAGTTCATGGTCAATGTGTGT
 GGAGACACCATAATAAACTCCTTTTTTGTGTTGCTAACTAAAACCACTAAAATCTAA
 AAACAAAACATTTTTGCACAAGAATTATTTATTATTCAATAAAGATGTTTAAATGGG
 GGAAGTTGAAGTTCATTGATAGTCTCATAAATCTTAAATGTATTTAACTGCTTTTTA
 CGTTTTTTATTATTAATTACTCTTGCTGTCATTATTATCATCATCATTATCGTCATCAT

CATCACTAATGCTTTTCACCATACACAAATGTAGGCAGAGAAGTGTGAATCCACTTAG
TGAGGCAATCTTGGAGAGGGAAAGGAAGCGGATGCGGGGCAGAGGCACACAGGAG
GACAGTGAGAGGGGAAATGAACAAGAAAAAATGTGGACACATGCACAAAAATTCCA
TAGTCCACTACATTACTTTGTATTCTAATATTAAGAAAAATAATAAACCCATTTCTGTG
CACTTATCACCCAGGCTCAACAGTTATCTTGGCCACAGATCCTGTCTCACTGCATCCT
GTCCACCTGAGTCCACTTAGCGTTCTGAATCCAATCCAGGGCATGATGCCTTACTCCT
ACACAGAACTAAAGATTAAAGAGAGTTTAAAAGTAACCATGACATCTCTCTGTTCCT
TTAGCGATAAGTTCCTAATATTTATGGCTGCTTGTGTATGTTCTAATTTCTCTAATATT
GTCACATTTAGTTGGCAACTACTTTGTTTGAATTGAGTTGGAGTTAAGGTCCCATAG
GATTAATCTCAACATATTTCTATATTTATAAACTTTTCTCTCTTTGTGAAAGTTCCTT
GAGAAAACAAATATGCCCATATCTTTCTTTACAGGTCTTAAAAATGAGCTATAACAA
AGTCCAAATAATTGAGAAGGATACTTTGTATGGACTCAGGAGCTTGACCCGGTTGCA
CCTGGATCACAAACACATTGAGTTTATCAACCCCGAGGCGTTTTACGGACTCACCTT
GCTCCGCTTGGTACATCTAGAAGGAAACCGGCTGACAAAGCTCCATCCAGACACAT
TTGTCTCTTTGAGCTATCTCCAGATATTTAAAACCTCCTTCATTAAGNACCTGTACTT
GTATGATAACTTCATTGACCTCCCTCCCAAAGAAATGGTCTCCTCTATGCCAAACC
TAGAAAGCCTTTACTTGCATGGAAACCCATGGACCTGTGACTGCCATTTAAAGTGGT
TGTCCGAGTGGATGCAGGGAAACCCAGGTAACATCTTGTGTTTGTGTTTCTTTTTTT
ATARKACGTATTTTCCTCAATTTCAATTTAGAATGATATCCCAAAGTCCCCCATAACC
TCCCCCCCCACTTCCCTACCTACCCATTCCCATTTTTTTGGCCCTGGCATTCCCCTGTACT
GGGGCATATAAAGTTTGCCTGTCCAATGGACCTCTCTTTCCAGTGATGGCCAACTAG
GCCATCTTTTGATACATATGCAGCTAGAGTCAAGAGCTCTGGGGTACTGGTTAGTTC
ATAATGTTGTTGCACCTACAGGGTTGAA (SEQ ID NO:20)

Figure 51

MPKRAHWGALS VV LILL WGHPR VALACPHPCACYVPSEVHCTFRSLASVPAGIARHVE
 RINLGFNSIQALSETSFAGLTKELELLMIHGNEIPSIPDGALRDLSSLQVFKFSYNKLRVITG
 QTLQGLSNLMRLHIDHNKIEFIHPQAFNGLTSLRLLHLEGNLLHQLHPSTFSTFTFLDYFR
 LSTIRHLYLAENMVRTLPA SMLRNMPLENLYLQGNPWTCDCEMRWFLWD AKSRGIL
 KCKKDKA YEGGQLCAMCFSPKKLYKHEIHKLKDMTCLKPSIESPLRQNRSR SIEEEQEQ
 EEDGGSQ LILEKFQLPQWSISLNM TDEHGNMVNLVCDIKKPM DVYKIHNLQTDPPDIDIN
 ATVALDFECPMTRENYEKLWKLIA YYSEVPVKLHRELMLS KDP RVSYQYRQDADEEAL
 YYTG VRAQILAEP EWVMQPSIDIQLNRRQSTAKKVLLSYYTQYSQTISTKDRQARGRS
 WVMIEPSGA VQRDQTVLEGGPCQLSCNVKASESPSIFWVLPDGSILKAPMDDPD SKFSIL
 SSGWLRIKSMEPSDSGLYQCIAQVRDEMDRMVYRVLVQSPSTQPAEKDVTIGKNPGES
 VTLPCNALAIPEAHL SWILPNRRIINDLANTSHVYMLPNGTLSIPKVQVSDSGYYRCVAV
 NQQGADHFTV GITVT KKGSGLP SKRGRRPGAKALSRVREDIVEDEGGSGMGDEENTSR
 RLLHPKDQEVFLKTKDDAINGDKKAKKGRRKLKLWKHSEKEPETNVAEGRRVFESRRR
 INMANKQINPERWADILAKVRGKNLPKGTEVPPLIKTTSPPSLSLEVTPPFPAVSPPSASP
 VQTVTSAEESSADVPLL GEEHVLTGISSASMGLEHNHNGVILVEPEVTSTPLEEVVDDL
 SEKTEEITSTEGDLKGTAAPT LISEPYEPSPTLHTLDTVYEKP THEETATEGWSAADVGSS
 PEPTSSEYEPPLDAVSLAESEPMQYFDPDLET KSQPDDEKMKEDTFAHLTPTPTIWVND S
 STSQLFEDSTIGEPGVPGQSHLQGLTDNIHLVKSSLSTQD TLLIKKG MKEMSQT LQGGNM
 LEGDP THSRSSSESEGQESKSITLPDSTLGIMSSMSPVKKPAETT VGTLLDKD TTTVTTPR
 QKVAPSSTMSTHPSRRRPNGRRLRP NKFRHRHKQTPPTTFAPSETFSTQPTQAPDIKISS
 QVSSLVPTAWVDNTVNTPKQLEMEKNAEPTSKGT PRRKHGKRPNKHRYTPSTVSSRA
 SGSKPSPSPENKH RNIVTPSSETILLPRTVSLKTEGPYDSLDMYTTTRKIYSSYPKVQETLP
 VTYKPTSDGKEIKDDVATNVDKHKSDILVTGESITNAIPTSRSLVSTMGEFKEESSPVGFP
 GTPTWNPSRTAQPGRLQTDIPVTTSGENLTD PPLLKELEDVDFTSEFLSSLTVSTPFHQEE
 AGSSTTLSSIKVEVASSQAETTTLDQDHLETTVA ILLSETRPQNHTPTAARMKEPASSPS
 TILMSLGQTTTTKPALPSPRISQASRDSKENVFLNYVGNPETEATPVNNEG TQHMSGPNE
 LSTPSSDRDAFNLSTKLELEKQVFGSRSLPRGPDSQRQDGRVHASHQLTRVPAKPILPTA
 TVRLPEMSTQSASRYFVTSQSPRHWTNKPEITTYPSGALPENKQFTTPRLSSTTIPLPLHM
 SKPSIPSKFTDRRTDQFNGYSKVFGNNNIPEARNPVGKPPSPRIPHYSNGRLPFFT NKTL SF
 PQLGVTRRPQIPTSPAPVMRERKVIPGSYNRIHSHSTFHLDFGPPAPLLHTPQT TGSPSTN
 LQNIPMV SSTQSSISFITSSVQSSGSFHQSSSKFFAGGPPASKFWSLGEKPQILTKSPQTVSV
 TAETDTVFPCEATGKPKPFVTWTKVSTGALMTPNTRIQRFEVLKNGTLVIRKVQVQDRG
 QYMCTASNLHGLDRMVVLLSVTVQQPQILASHYQDVTVYLGDTIAMECLAKGTPAPQI
 SWIFPDRRVWQTVSPVESRITLHENRTLSIKEASFSDRGVYKCVASNAAGADSLAIRLHV
 AALPPVIHQEKLENISLPPGLSIHIHCTAKAAPLPSVRWVLGDGTQIRPSQFLHGNLFVFP
 NGTLYIRNLAPKDSGRYECVAANLVGSARRTVQLNVQRAAANARITGTS PRRTDVRYG
 GTLKLDCSASGDPWPRILWRLPSKRMIDALFSFDSRIKVFANGTLVVKSVTDK

DAGDYLCVARNKVGDDYVVLKVDVVMKPAKIEHKEENDHKVIFYGGDLKVDCVATGL
PNPEISWSLPDGSLVNSFMQSDDSGGRTKRYVVFNNGTLYFNEVGMREEGDYTCFAEN
QVGKDEMVRVRVKVVTAPATIRNKTYLAVQVPYGDVVTVACEAKGEPMPKVTWLSPTN
KVIPTSSEKYQIYQDGTLLIQKAQRSDSGNYTCLVRNSAGEDRKTVWIHVNVPKING
NPNPITTVREIAAGGSRKLDCKAEGIPTPRVLWAFPEGVVLPAPIYGNRITVHGNGSLDI
RSLRKSDSVQLVCMARNEGGEARLIVQLTVLEPMEKPIFHDPISEKITAMAGHTISLNCS
AAGTPTPSLVWVLPNGTDLQSGQQLQRFYHKADGMLHISGLSSVDAGAYRCVARNA
GHTERLVSLKVGLKPEANKQYHNLVSIINGETLKLPPGAGQGRFSWTLPNGMHLEG
PQTLGRVSLDNGTLTVREASVFDRGTYVCRMETEYGPSVTSIPVIVIAYPFRITSEPTVI
YTRPGNTVKLNCMAMGIPKADITWELPDKSHLKAGVQARLYGNRFLHPQGSLLTIQHAT
QRDAGFYKCMANKILGSDSKTTYIHVF (SEQ ID NO: 21)

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

Figure 52

ATGCCCAAGC GCGCGCACTG GGGGGCCCTC TCCGTGGTGC TGATCCTGCT
 TTGGGGCCAT CCGCGAGTGG CGCTGGCCTG CCCGCATCCT TGTGCCTGCT
 ACGTCCCCAG CGAGGTCCAC TGCACGTTCC GATCCCTGGC TTCCGTGCCC
 GCTGGCATTG CTAGACACGT GGAAAGAATC AATTTGGGGT TTAATAGCAT
 ACAGGCCCTG TCAGAAACCT CATTTGCAGG ACTGACCAAG TTGGAGCTAC
 TTATGATTCA CGGCAATGAG ATCCCAAGCA TCCCGATGG AGCTTTAAGA
 GACCTCAGCT CTCTTCAGGT TTTCAAGTTC AGCTACAACA AGCTGAGAGT
 GATCACAGGA CAGACCCTCC AGGGTCTCTC TAACTTAATG AGGCTGCACA
 TTGACCACAA CAAGATCGAG TTTATCCACC CTCAAGCTTT CAACGGCTTA
 ACGTCTCTGA GGCTACTCCA TTTGGAAGGA AATCTCCTCC ACCAGCTGCA
 CCCCAGCACC TTCTCCACGT TCACATTTTT GGATTATTTT AGACTCTCCA
 CCATAAGGCA CCTCTACTTA GCAGAGAACA TGGTTAGAAC TCTTCCTGCC
 AGCATGCTTC GGAACATGCC GCTTCTGGAG AATCTTTACT TGCAGGGAAA
 TCCGTGGACC TGCGATTGTG AGATGAGATG GTTTTTGGAA TGGGATGCAA
 AATCCAGAGG AATTCTGAAG TGTA AAAAGG ACAAAGCTTA TGAAGGCGGT
 CAGTTGTGTG CAATGTGCTT CAGTCCAAAG AAGTTGTACA AACATGAGAT
 ACACAAGCTG AAGGACATGA CTTGTCTGAA GCCTTCAATA GAGTCCCCTC
 TGAGACAGAA CAGGAGCAGG AGTATTGAGG AGGAGCAAGA ACAGGAAGAG
 GATGGTGGCA GCCAGCTCAT CCTGGAGAAA TTCCAACCTGC CCCAGTGGAG
 CATCTCTTTG AATATGACCG ACGAGCACGG GAACATGGTG AACTTGGTCT
 GTGACATCAA GAAACCAATG GATGTGTACA AGATTCACTT GAACCAAACG
 GATCCTCCAG ATATTGACAT AAATGCAACA GTTGCCTTGG ACTTTGAGTG
 TCCAATGACC CGAGAAAACCT ATGAAAAGCT ATGGAAATTG ATAGCATACT
 ACAGTGAAGT TCCCGTGAAG CTACACAGAG AGCTCATGCT CAGCAAAGAC
 CCCAGAGTCA GCTACCAGTA CAGGCAGGAT GCTGATGAGG AAGCTCTTTA
 CTACACAGGT GTGAGAGCCC AGATTCTTGC AGAACCAGAA TGGGTCATGC
 AGCCATCCAT AGATATCCAG CTGAACCGAC GTCAGAGTAC GGCCAAGAAG
 GTGCTACTTT CTTACTACAC CCAGTATTCT CAAACAATAT CCACCAAAGA
 TACAAGGCAG GCTCGGGGCA GAAGCTGGGT AATGATTGAG CCTAGTGGAG
 CTGTGCAAAG AGATCAGACT GTCCTGGAAG GGGGTCCATG CCAGTTGAGC
 TGCAACGTGA AAGCTTCTGA GAGTCCATCT ATCTTCTGGG TGCTTCCAGA
 TGGCTCCATC CTGAAAGCGC CCATGGATGA CCCAGACAGC AAGTTCTCCA
 TTCTCAGCAG TGGCTGGCTG AGGATCAAGT CCATGGAGCC ATCTGACTCA
 GGCTTGTACC AGTGCATTGC TCAAGTGAGG GATGAAATGG ACCGCATGGT
 ATATAGGGTA CTTGTGCAGT CTCCCTCCAC TCAGCCAGCC GAGAAAGACA
 CAGTGACAAT TGGCAAGAAC CCAGGGGAGT CGGTGACATT GCCTTGCAAT
 GCTTTAGCAA TACCCGAAGC CCACCTTAGC TGGATTCTTC CAAACAGAAG
 GATAATTAAT GATTTGGCTA ACACATCACA TGTATACATG TTGCCAAATG
 GAACTCTTTC CATCCCAAAG GTCCAAGTCA GTGATAGTGG

TTA CTACAGA TGTGTGGCTG TCAACCAGCA AGGGGCAGAC CATTTTACGG
 TGGAATCAC AGTGACCAAG AAAGGGTCTG GCTTGCCATC CAAAAGAGGC
 AGACGCCAG GTGCAAAGGC TCTTTCCAGA GTCAGAGAAG ACATCGTGGA
 GGATGAAGGG GGCTCGGGCA TGGGAGATGA AGAGAACACT TCAAGGAGAC
 TTCTGCATCC AAAGGACCAA GAGGTGTTCC TCAAAACAAA GGATGATGCC
 ATCAATGGAG ACAAGAAAGC CAAGAAAGGG AGAAGAAAGC TGAAACTCTG
 GAAGCATTCG GAAAAAGAAC CAGAGACCAA TGTTGCAGAA GGTCGCAGAG
 TGTTTGAATC TAGACGAAGG ATAAACATGG CAAACAAACA GATTAATCCG
 GAGCGCTGGG CTGATATTTT AGCCAAAGTC CGTGGGAAAA ATCTCCCTAA
 GGGCACAGAA GTACCCCGT TGATTAAC CACAAGTCCT CCATCCTTGA
 GCCTAGAAGT CACACCACCT TTTCTGCTG TTTCTCCCC CTCAGCATCT
 CCTGTGCAGA CAGTAACCAG TGCTGAAGAA TCCTCAGCAG ATGTACCTCT
 ACTTGGTGAA GAAGAGCACG TTTTGGGTAC CATTTCTCA GCCAGCATGG
 GGCTAGAACA CAACCACAAT GGAGTTATC TTGTTGAACC TGAAGTAACA
 AGCACACCTC TGGAGGAAGT TGTTGATGAC CTTTCTGAGA AGACTGAGGA
 GATAACTTCC ACTGAAGGAG ACCTGAAGGG GACAGCAGCC CCTACACTTA
 TATCTGAGCC TTATGAACCA TCTCCTACTC TGCACACATT AGACACAGTC
 TATGAAAAGC CCACCCATGA AGAGACGGCA ACAGAGGGTT GGTCTGCAGC
 AGATGTTGGA TCGTCACCAG AGCCACATC CAGTGAGTAT GAGCCTCCAT
 TGGATGCTGT CTCCTTGGCT GAGTCTGAGC CCATGCAATA CTTTGACCCA
 GATTTGGAGA CTAAGTCACA ACCAGATGAG GATAAGATGA AAGAAGACAC
 CTTTGCACAC CTTACTCCAA CCCCACCAT CTGGGTTAAT GACTCCAGTA
 CATCACAGTT ATTTGAGGAT TCTACTATAG GGGAACCAGG TGTCACAGGC
 CAATCACATC TACAAGGACT GACAGACAAC ATCCACCTTG TGAAAAGTAG
 TCTAAGCACT CAAGACACCT TACTGATTAA AAAGGGTATG AAAGAGATGT
 CTCAGACACT ACAGGGAGGA AATATGCTAG AGGGAGACCC CACACACTCC
 AGAAGTTCTG AGAGTGAGGG CCAAGAGAGC AAATCCATCA CTTTGCCTGA
 CTCCACACTG GGTATAATGA GCAGTATGTC TCCAGTTAAG AAGCCTGCGG
 AAACCACAGT TGGTACCCTC CTAGACAAAG ACACCACAAC AGTAACAACA
 ACACCAAGGC AAAAAGTTGC TCCGTCATCC ACCATGAGCA CTCACCCTTC
 TCGAAGGAGA CCAACGGGA GAAGGAGATT ACGCCCCAAC AAATTCCGCC
 ACCGGCACAA GCAAACCCCA CCCACAACCTT TTGCCCCATC AGAGACTTTT
 TCTACTCAAC CAACTCAAGC ACCTGACATT AAGATTTCAA GTCAAGTGGA
 GAGTTCTCTG GTTCCTACAG CTGGGTGGA TAACACAGTT AATACCCCA
 AACAGTTGGA AATGGAGAAG AATGCAGAAC CCACATCCAA GGGAACACCA
 CGGAGAAAAC ACGGGAAGAG GCCAAACAAA CATCGATATA CCCCTTCTAC
 AGTGAGCTCA AGAGCGTCCG GATCCAAGCC CAGCCCTTCT CCAGAAAATA
 AACATAGAAA CATTGTTACT CCCAGTTCAG AAATATACT TTTGCCTAGA
 ACTGTTTCTC TGAAAACCTGA GGGCCCTTAT GATTCCTTAG ATTACATGAC
 AACCACCAGA AAAATATATT CATCTTACCC TAAAGTCCAA GAGACACTTC
 CAGTCACATA

TAAACCCACA TCAGATGGAA AAGAAATTAA GGATGATGTT GCCACAAATG
 TTGACAAACA TAAAAGTGAC ATTTTAGTCA CTGGTGAATC AATTACTAAT
 GCCATACCAA CTTCTCGCTC CTTGGTCTCC ACTATGGGAG AATTTAAGGA
 AGAATCCTCT CCTGTAGGCT TTCCAGGAAC TCCAACCTGG AATCCCTCAA
 GGACGGCCCA GCCTGGGAGG CTACAGACAG ACATACCTGT TACCACTTCT
 GGGGAAAATC TTACAGACCC TCCCCTTCTT AAAGAGCTTG AGGATGTGGA
 TTCACTTCC GAGTTTTTGT CCTCTTTGAC AGTCTCCACA CCATTTCACC
 AGGAAGAAGC TGGTCTTCC ACAACTCTCT CAAGCATAAA AGTGGAGGTG
 GCTTCAAGTC AGGCAGAAAC CACCACCCTT GATCAAGATC ATCTTGAAAC
 CACTGTGGCT ATTCTCCTTT CTGAAACTAG ACCACAGAAT CACACCCCTA
 CTGCTGCCCCG GATGAAGGAG CCAGCATCCT CGTCCCCATC CACAATTCTC
 ATGTCTTTGG GACAAACCAC CACCACTAAG CCAGCACTTC CCAGTCCAAG
 AATATCTCAA GCATCTAGAG ATTCCAAGGA AAATGTTTTT TTGAATTATG
 TGGGGAATCC AGAAACAGAA GCAACCCCAG TCAACAATGA AGGAACACAG
 CATATGTCAG GGCCAAATGA ATTATCAACA CCCTCTTCCG ACCGGGATGC
 ATTTAACTTG TCTACAAAGC TGGAATTGGA AAAGCAAGTA TTTGGTAGTA
 GGAGTCTACC ACGTGGCCCA GATAGCCAAC GCCAGGATGG AAGAGTTCAT
 GCTTCTCATC AACTAACCAG AGTCCCTGCC AAACCCATCC TACCAACAGC
 AACAGTGAGG CTACCTGAAA TGTCCACACA AAGCGCTTCC AGATACTTTG
 TAACTTCCCA GTCACCTCGT CACTGGACCA ACAAACCGGA AATAACTACA
 TATCCTTCTG GGGCTTTGCC AGAGAACAAA CAGTTTACAA CTCCAAGATT
 ATCAAGTACA ACAATTCCTC TCCATTGCA CATGTCCAAA CCCAGCATTC
 CTAGTAAGTT TACTGACCGA AGAACTGACC AATTCAATGG TTA CTCCAAA
 GTGTTTGGA ATAACAACAT CCCTGAGGCA AGAAACCCAG TTGGAAAGCC
 TCCCAGTCCA AGAATTCCTC ATTATTCCAA TGGAAGACTC CCTTTCTTTA
 CCAACAAGAC TCTTTCTTTT CCACAGTTGG GAGTCACCCG GAGACCCCG
 ATACCCACTT CTCCTGCCCC AGTAATGAGA GAGAGAAAAG TTATTCCAGG
 TTCCTACAAC AGGATACATT CCCATAGCAC CTTCATCTG GACTTTGGCC
 CTCCGGCACC TCCGTTGTTG CACACTCCGC AGACCACGGG ATCACCTCA
 ACTAACTTAC AGAATATCCC TATGGTCTCT TCCACCCAGA GTTCTATCTC
 CTTTATAACA TCTTCTGTCC AGTCCTCAGG AAGCTTCCAC CAGAGCAGCT
 CAAAGTTCTT TGCAGGAGGA CCTCCTGCAT CCAAATTCTG GTCTCTTGGG
 GAAAAGCCCC AAATCCTCAC CAAGTCCCCA CAGACTGTGT CCGTCACCGC
 TGAGACAGAC ACTGTGTTCC CCTGTGAGGC AACAGGAAAA CCAAAGCCTT
 TCGTTACTTG GACAAAGGTT TCCACAGGAG CTCTTATGAC TCCGAATACC
 AGGATACAAC GGT TTGAGGT TCTCAAGAAC GGTACCTTAG TGATACGGAA
 GGTTCAGTA CAAGATCGAG GCCAGTATAT GTGCACCGCC AGCAACCTGC
 ACGGCCTGGA CAGGATGGTG GTCTTGCTTT CGGTCACCGT GCAGCAACCT
 CAAATCCTAG CCTCCCACTA CCAGGACGTC ACTGTCTACC TGGGAGACAC
 CATTGCAATG GAGTGTCTGG CCAAAGGGAC CCCAGCCCCC CAAATTTCCT
 GGATCTTCCC TGACAGGAGG GTGTGGCAAA CTGTGTCCCC

CGTGGAGAGC CGCATCACCC TGCACGAAAA CCGGACCCTT TCCATCAAGG
AGGCGTCCTT CTCAGACAGA GCGTCTATA AGTGCGTGGC CAGCAATGCA
GCCGGGGCGG ACAGCCTGGC CATCCGCCTG CACGTGGCGG CACTGCCCCC
CGTTATCCAC CAGGAGAAGC TGGAGAACAT CTCGCTGCCC CCGGGGCTCA
GCATTCACAT TCACTGCACT GCCAAGGCTG CGCCCCTGCC CAGCGTGCGC
TGGGTGCTCG GGGACGGTAC CCAGATCCGC CCCTCGCAGT TCCTCCACGG
GAACTTGTTT GTTTTCCCCA ACGGGACGCT CTACATCCGC AACCTCGCGC
CCAAGGACAG CGGGCGCTAT GAGTGCCTGG CCGCCAACCT GGTAGGCTCC
GCGCGCAGGA CGGTGCAGCT GAACGTGCAG CGTGCAGCAG CCAACGCGCG
CATCACGGGC ACCTCCCCGC GGAGGACGGA CGTCAGGTAC GGAGGAACCC
TCAAGCTGGA CTGCAGCGCC TCGGGGGACC CCTGGCCGCG CATCCTCTGG
AGGCTGCCGT CCAAGAGGAT GATCGACGCG CTCTTCAGTT TTGATAGCAG
AATCAAGGTG TTTGCCAATG GGACCCTGGT GGTGAAATCA GTGACGGACA
AAGATGCCGG AGATTACCTG TGCCTAGCTC GAAATAAGGT TGGTGATGAC
TACGTGGTGC TCAAAGTGGA TGTGGTGATG AAACCGGCCA AGATTGAACA
CAAGGAGGAG AACGACCACA AAGTCTTCTA CGGGGGTGAC CTGAAAGTGG
ACTGTGTGGC CACCGGGCTT CCCAATCCCG AGATCTCCTG GAGCCTCCCA
GACGGGAGTC TGGTGAAGTC CTTCATGCAG TCGGATGACA GCGGTGGACG
CACCAAGCGC TATGTCGTCT TCAACAATGG GACACTCTAC TTAAACGAAG
TGGGGATGAG GGAGGAAGGA GACTACACCT GCTTTGCTGA AAATCAGGTC
GGGAAGGACG AGATGAGAGT CAGAGTCAAG GTGGTGACAG CGCCCGCCAC
CATCCGGAAC AAGACTTACT TGGCGGTTCA GGTGCCCTAT GGAGACGTGG
TCACTGTAGC CTGTGAGGCC AAAGGAGAAC CCATGCCCAA GGTGACTTGG
TTGTCCCCAA CCAACAAGGT GATCCCCACC TCCTCTGAGA AGTATCAGAT
ATACCAAGAT GGCCTCTCC TTATTAGAA AGCCAGCGT TCTGACAGCG
GCAACTACAC CTGCCTGGTC AGGAACAGCG CGGGAGAGGA TAGGAAGACG
GTGTGGATTG ACGTCAACGT CCAGCCACCC AAGATCAACG GTAACCCCAA
CCCCATCACC ACCGTGCGGG AGATAGCAGC CGGGGGCAGT CGGAAACTGA
TTGACTGCAA AGCTGAAGGC ATCCCCACCC CGAGGGTGTT ATGGGCTTTT
CCCGAGGGTG TGGTTCTGCC AGCTCCATAC TATGGAAACC GGATCACTGT
CCATGGCAAC GGTTCCTGG ACATCAGGAG TTTGAGGAAG AGCGACTCCG
TCCAGCTGGT ATGCATGGCA CGCAACGAGG GAGGGGAGGC GAGGTTGATC
GTGCAGCTCA CTGTCCTGGA GCCCATGGAG AAACCCATCT TCCACGACCC
GATCAGCGAG AAGATCACGG CCATGGCGGG CCACACCATC AGCCTCAACT
GCTCTGCCGC GGGGACCCCG ACACCCAGCC TGGTGTGGGT CCTTCCCAAT
GGCACCGATC TGCAGAGTGG ACAGCAGCTG CAGCGCTTCT ACCACAAGGC
TGACGGCATG CTACACATTA GCGGTCTCTC CTCGGTGGAC GCTGGGGCCT
ACCGCTGCGT GGCCCGCAAT GCCGCTGGCC ACACGGAGAG GCTGGTCTCC
CTGAAGGTGG GACTGAAGCC AGAAGCAAAC AAGCAGTATC ATAACCTGGT
CAGCATCATC AATGGTGAGA CCCTGAAGCT CCCCTGCACC CCTCCCGGGG
CTGGGCAGGG ACGTTTCTCC TGGACGCTCC CCAATGGCAT GCATCTGGAG

GGCCCCAAA CCCTGGGACG CGTTTCTCTT CTGGACAATG GCACCCTCAC
GGTTCGTGAG GCCTCGGTGT TTGACAGGGG TACCTATGTA TGCAGGATGG
AGACGGAGTA CGGCCCTTCG GTCACCAGCA TCCCCGTGAT TGTGATCGCC
TATCCTCCCC GGATCACCAG CGAGCCCACC CCGGTCATCT ACACCCGGCC
CGGGAACACC GTGAAACTGA ACTGCATGGC TATGGGGATT CCCAAAGCTG
ACATCACGTG GGAGTTACCG GATAAGTCGC ATCTGAAGGC AGGGGTTCAG
GCTCGTCTGT ATGGAAACAG ATTTCTTCAC CCCAGGGGAT CACTGACCAT
CCAGCATGCC ACACAGAGAG ATGCCGGCTT CTACAAGTGC ATGGCAAAAA
ACATTCTCGG CAGTGACTCC AAAACAACCTT ACATCCACGT CTTCTGAAAT
GTGGATTCCA GAATGATTGC TTAGGAACTG ACAACAAAGC GGGGTTCAG
AGGGAAGCCA GGTGTTGGGAA TAGGAGCTCT TAAATAATGT GTCACAGTGC
ATGGTGGCCT CTGGTGGGTT TCAAGTTGAG GTTGATCTTG ATCTACAATT
GTTGGGAAAA GGAAGCAATG CAGACACGAG AAGGAGGGCT CAGCCTTGCT
GAGACACTTT CTTTGTGTT TACATCATGC CAGGGGCTTC ATTCAGGGTG
TCTGTGCTCT GACTGCAATT TTTCTTCTTT TGCAAATGCC ACTCGACTGC
CTTCATAAGC GTCCATAGGA TATCTGAGGA ACATTATCA AAAATAAGCC
ATAGACATGA ACAACACCTC ACTACCCCAT TGAAGACGCA TCACCTAGTT
AACCTGCTGC AGTTTTTACA TGATAGACTT TGTTCCAGAT TGACAAGTCA
TCTTTCAGTT ATTTCTCTG TCACTTCAA ACTCCAGCTT GCCCAATAAG
GATTTAGAAC CAGAGTGACT GATATATATA TATATATTTT AATTCAGAGT
TACATACATA CAGCTACCAT TTTATATGAA AAAAGAAAAA CATTTCTTCC
TGGAACCTCAC TTTTATATA ATGTTTTATA TATATATTTT TTCCTTTCAA
ATCAGACGAT GAGACTAGAA GGAGAAATAC TTTCTGTCTT ATTAAAATTA
ATAAATTATT GGTCTTTACA AGACTTGGAT ACATTACAGC AGACATGGAA
ATATAATTTT AAAAAATTC TCTCCAACCT CCTTCAAATT CAGTCACCAC
TGTTATATTA CCTTCTCCAG GAACCCTCCA GTGGGGAAGG CTGCGATATT
AGATTTCTT GTATGCAAAG TTTTGTGTA AAGCTGTGCT CAGAGGAGGT
GAGAGGAGAG GAAGGAGAAA ACTGCATCAT AACTTTACAG AATTGAATCT
AGAGTCTTCC CCGAAAAGCC CAGAACTTC TCTGCAGTAT CTGGCTTGTC
CATCTGGTCT AAGGTGGCTG CTTCTTCCCC AGCCATGAGT CAGTTTGTGC
CCATGAATAA TACACGACCT GTTATTTCCA TGACTGCTTT ACTGTATTTT
TAAGGTCAAT ATACTGTACA TTTGATAATA AAATAATATT CTCCCCAAAA AAAAA

Human OCP: nucleotide sequence of ORF

ATGAAGGTAAAAGGCAGAGGAATCACCTGCTTGCTGGTCTCCTTTGCTGTGATCTGC
CTGGTCGCCACCCCTGGGGGCAAGGCCTGTCTCGCCGCTGTGCCTGTTATATGCCT
ACGGAGGTACACTGCACATTTTCGGTACCTGACTTCCATCCCAGACAGCATCCCGCCC
AATGTGGAACGCATCAATTTAGGATACAACAGCTTGGTTAGATTGATGGAAACAGAT
TTTTCTGGCCTGACCAAACCTGGAGTTACTCATGCTTCACAGCAATGGCATTACACA
ATCCCTGACAAGACCTTCTCAGATTTGCAGGCCTTGCAGGTCTTAAAAATGAGCTAT
AATAAAGTCCGAAACCTTCAGAAAGATACTTTTTATGGCCTCAGGAGCTTGACACGA
TTGCACATGGACCACAACAATATTGAGTTTATAAACCCAGAGGTTTTTTATGGGCTC
AACTTTCTCCGCCTGGTGCCTTGGAAAGGAAATCAGCTCACTAAGCTCCACCCAGAT
ACATTTGTCTCTTTGAGCTACCTCCAGATATTTAAATCTCTTTCATTAAGTTCCTA
TACTTGTCTGATAACTTCCTGACCTCCCTCCCTCAAGAGATGGTCTCCTATATGCCT
GACCTAGACAGCCTTTACCTGCATGGAAACCCATGGACCTGTGATTGCCATTTAAAG
TGGTTGTCTGACTGGATACAGGAGAAGCCAGATGTAATAAAATGCAAAAAGATAGA
AGTCCCTCTAGTGCTCAGCAGTGTCCACTTTGCATGAACCCTAGGACTTCTAAAGGC
AAGCCGTTAGCTATGGTCTCAGCTGCAGCTTTCAGTGTGCCAAGCCAACCATTGAC
TCATCCCTGAAATCAAAGAGCCTGACTATTCTGGAAGACAGTAGTTCTGCTTTCATC
TCTCCCAAGGTTTCATGGCACCCTTTGGCTCCCTCACTTTGAATATGACAGATCAG
TCTGGAAATGAAGCTAACATGGTCTGCAGTATTCAAAAGCCCTCAAGGACATCACCC
ATTGCATTCACTGAAGAAAATGACTACATCGTGCTAAATACTTCATTTTCAACATTT
TTGGTGTGCAACATAGATTACGGTCACATTCAGCCAGTGTGGCAAATTTTGGCTTTG
TACAGTGATTCTCCTCTGATACTAGAAAGGAGCCACTTGCTTAGTGAAACACCGCAG
CTCTATTACAAATATAAACAGGTGGCTCCTAAGCCTGAAGACATTTTTACCAACATA
GAGGCAGATCTCAGAGCAGATCCCTCTTGGTTAATGCAAGACCAAATTTCTTGCAG
CTGAACAGAACTGCCACCACATTCAGTACATTACAGATCCAGTACTCCAGTGATGCT
CAAATCACTTTACCAAGAGCAGAGATGAGGCCAGTGAAACACAAATGGACTATGATT
TCAAGGGATAACAATACTAAGCTGGAACATACTGTCTTGGTAGGTGGAACCGTTGGC
CTGAAGTAAAGTGAGAGCCCCCTTATGTGCTGAGGATGGACGGATCCTAATAGACAAA
AGTGGAATAATTGGAACCTCAGATGGCTGATAGTTTTTGACACAGGCGTATATCACTGT
ATAAGCAGCAATTATGATGATGCAGATATTCTCACCTATAGGATAACTGTGGTAGAA
CCTTTGGTCGAAGCCTATCAGGAAAATGGGATTTCATCACACAGTTTTTCATTGGTGAA
ACACTTGATCTTCCATGCCATTCTACTGGTATCCCAGATGCCTCTATTAGCTGGGTT
ATTCCAGGAAACAATGTGCTCTATCAGTCATCAAGAGACAAGAAAGTTCTAAACAAT
GGCACATTAAGAATATTACAGGTCACCCCGAAAGACCAAGGTTATTATCGCTGTGTG
GCAGCCAACCCATCAGGGGTTGATTTTTTTGATTTTCCAAGTTTCAGTCAAGATGAAA
GGACAAAGGCCCTTGGAGCATGATGGAGAAACAGAGGGATCTGGACTTGATGAGTCC
AATCCTATTGCTCATCTTAAGGAGCCACCAGGTGCACAACTCCGTACATCTGCTCTG
ATGGAGGCTGAGGTTGGAAAACACACCTCAAGCACAAAGTAAGAGGCACAACTATCGG
GAATTAACACTCCAGCGACGTGGAGATTCAACACATCGACGTTTTAGGGAGAATAGG
AGGCATTTCCCTCCCTCTGCTAGGAGAATTGACCCACAACATTGGGCGGCACTGTTG
GAGAAAGCTAAAAAGAATGCTATGCCAGACAAGCGAGAAAATACCACAGTGAGCCCA
CCCCAGTGGTCACCCAACCTCCCAAACATACCTGGTGAAGAAGACGATTCTCAGGC
ATGCTCGCTCTACATGAGGAATTTATGGTCCCGGCCACTAAAGCTTTGAACCTTCCA
GCAAGGACAGTGACTGCTGACTCCAGAACAATATCTGATAGTCCTATGACAAACATA
AATTATGGCACAGAATTCTCTCCTGTTGTGAATTCACAAATACTACCACCTGAAGAA
CCCACAGATTTCAAACGTCTACTGCTATTAAACTACAGCCATGTCAAAGAATATA

AACCCCAACCATGTCAAGCCAAATACAAGGCACAACCAATCAACATTCCATCCACTGTC
TTTCCACTGCTACTTTGGAGCAACTGAATTTTCAGGACTCTGACCAGATGGGAAGAGGA
AGAGAGCATTTCCAAAGTAGACCCCCAATAACAGTAAGGACTATGATCAAAGATGTC
AATGTCAAAATGCTTAGTAGCACCACCAACAACTATTATTAGAGTCAGTAAATACC
ACAAATAGTCATCAGACATCTGTAAGAGAAGTGAGTGAACCCAGGCACAATCACTTC
TATTCTCACACTACTCAAATACTTAGCACCTCCACGTTCCCTTCAGATCCACACACA
GCTGCTCATTCTCAGTTTCCGATCCCTAGAAATAGTACAGTTAACATCCCGCTGTTC
AGACGCTTTGGGAGGCAGAGGAAAATTGGCGGAAGGGGGCGGATTATCAGCCCATAT
AGAACTCCAGTTCTGCGACGGCATAGATACAGCATTTTTCAGGTCAACAACCAGAGGT
TCTTCTGAAAAAAGCACTACTGCATTCTCAGCCACAGTGCTCAATGTGACATGTCTG
TCCTGTCTTCCCAGGGAGAGGCTCACCCTGCCACAGCAGCATTGTCTTTTCCAAGT
GCTGCTCCCATCACCTTCCCCAAAGCTGACATTGCTAGAGTCCCATCAGAAGAGTCT
ACAACTCTAGTCCAGAATCCACTATTACTACTTGAGAACAAACCCAGTGTAGAGAAA
ACAACACCCACAATAAAATATTTTCAGGACTGAAATTTCCCAAGTGACTCCAACGGT
GCAGTCATGACATATGCTCCAACATCCATACCCATGGAAAAAACTCACAAAGTAAAC
GCCAGTTACCCACGTGTGTCTAGCACCAATGAAGCTAAAAGAGATTTCAGTGATTACA
TCGTCACTTTTCAGGTGCTATCACCAAGCCACCAATGACTATTATAGCCATTACAAGG
TTTTCAAGAAGGAAAATTCCCTGGCAACAGAACTTTGTAAATAACCATAACCCAAAA
GGCAGATTAAAGGAATCAACATAAAGTTAGTTTACAAAAAAGCACAGCTGTGATGCTT
CCTAAAACATCTCCTGCTTTACCACAGAGACAAAGTTCCCTTTTCCATTTTACCACA
CTTTCAACAAGTGTGATGCAAATTCATCTAATACCTTGACTACCGCTCACCACACT
ACGACCAAAAACACACAATCCTGGAAGTCTTCCAACAAAGAAGGAGCTTCCCTTCCCA
CCCCTTAACCTTATGCTTCTTAGTATTATAAGCAAAGACTCAAGTACAAAAAGCATC
ATATCAACGCAAACAGCAATACCAGCAACAACCTTACCTTCCCTGCATCTGTCTATC
ACTTATGAAACCCAAACAGAGAGATCTAGAGCACAAACAATACAAAGAGAACAGGAG
CCTCAAAAGAAGAACAGGACTGACCCAAACATCTCTCCAGACCAGAGTTCTGGCTTC
ACTACACCCACTGCTATGACACCTCCTGCTCTGGCATTCACTCATTCCCCACCAGAA
AACACAACCTGGGATTTCAAGCACAAATCAGTTTTTCATTCAAGAACTCTTAATCTGACA
GATGTGATTGAAGAACTAGCCCAAGCAAGTACTCAGACTTTGAAGAGCACAAATTGCT
TCTGAAACAACCTTTGTCCAGCAAATCACACCAGAGTACCACAACCTAGGAAAGCATCA
TTAGACACTCCCATACCACCATTCTTGAGCAGCAGTGCTACTCTAATGCCAGTTCCC
ATCTCCCCCTCCCTTTACTCAGAGAGCAGTTACTGACACACGTGGCGACTCCCATTTC
CGGCTTATGACAAATACAGTGGTCAAGCTGCACGAATCCTCAAGGCACAATCTCCAA
ATGCCAAGTTCACAATTGGAACCACTCACTTCATCTACCTCTAATCTGTTACATTCT
ACTCCCATGCCAGCACTAACAAACAGTTAAATCACAGAATTCCAAATTAACCTCCATCT
CCCTGGGCAGAATACCAATTTTGGCACAAACCATACTCAGACATTGCTGAAAAAGGC
AAAAAGCCAGAAGTAAGCATGTTGGCTACTACAGGCCTGTCCGAGGCCACCACTCTT
GTTTCAGATTGGGATGGACAGAAGAACACAAAGAAGAGTGACTTTGATAAGAAACCA
GTTCAAGAAGCAACAACCTTCCAACTCCTTCCCTTTGACTCTTTGTCTAGGTATATA
TTTGAAGAAGCCAGGATAGTTGGAGGAAAAGCTGCAAGTTTTTACTATTCCAGCTAAC
TCAGATGCCTTTCTTCCCTGTGAAGCTGTTGGAAATCCCCTGCCACCATTTCATTGG
ACCAGAGTTTCAGGACTTGATTTATCTAGAGGAAACCAGAATAGCAGGGTCCAGGTT
CTCCCCAATGGTACCCTGTCCATCCAGAGGGTGGAAATTCAGGACCGCGGACAGTAC
TTGTGTTCCGCATCCAATCTGTTTGGCACAGACCACCTTCATGTACCTTGTCTGTG
GTTTCCTATCCTCCCAGGATCCTGGAGAGACGTACCAAAGAGATCACAGTTCATTCC
GGAAGCACTGTGGAAGTGAAGTGCAGAGCAGAAGGTAGGCCAAGCCCTACAGTTACC
TGGATTCTTGCAAACCAAAACAGTTGTCTCAGAATCATCCCAGGGAAGTAGGCAGGCT
GTGGTGACGGTTGACGGAACATTGGTCCTCCACAATCTCAGTATTTATGACCGTGGC
TTTTACAAATGTGTGGCCAGCAACCCAGGTGGCCAGGATTCACTGCTGGTTAAAATA
CAAGTCATTGCAGACCACTGTTATTCTAGAGCAAAGGAGGCAAGTCATTGTAGGC
ACTTGGGGTGAAAGTTTAAAACCTGCCCTGTACTGCAAAAGGAACTCCTCAGCCCAGC

GTT TACTGGGTCCTCTCTGATGGCACTGAAGTGAAACCATTACAGTTTACCAATTCC
 AAGTTGTTCTTATTTTCAAATGGGACTTTGTATATAAGAAACCTAGCCTCTTCAGAC
 AGGGGCACTTATGAATGCATTGCTACCAGTTCCACTGGTTTCGGAGCGAAGAGTAGTA
 ATGCTTACAATGGAAGAGCGAGTGACCAGCCCCAGGATAGAAGCTGCATCCCAGAAA
 AGGACTGAAGTGAATTTTGGGGACAAATTACTACTGAACTGCTCAGCCACTGGGGAG
 CCCAAACCCCCAAATAATGTGGAGGTTACCATCCAAGGCTGTGGTCGACCAGTGGAGC
 TGGATCCACGCTACCCCTAATGGATCCCTGTTTATTGGATCAGTAACAGAAAAAGAC
 AGTGGTGTCTACTTGTGTGTGGCAAGAAACAAAATGGGGGATGATCTGATACTGATG
 CATGTTAGCCTAAGACTGAAACCTGCCAAAATTGACCACAAGCAGTATTTTAGAAAG
 CAAGTGCTCCATGGGAAAGATTTCCAAGTAGATTGCAAAGCTTCCGGCTCCCCAGTG
 CCAGAGATATCTTGGAGTTTGCCTGATGGAACCATGATCAACAATGCAATGCAAGCC
 GATGACAGTGGCCACAGGACTAGGAGATATACCCTTTTCAACAATGGAACCTTTATAC
 TTCAACAAAGTTGGGGTAGCGGAGGAAGGAGATTATACTTGCTATGCCCAGAACACC
 CTAGGGAAAGATGAAATGAAGGTCCACTTAACAGTTATAACAGCTGCTCCCCGGATA
 AGGCAGAGTAACAAAACCAACAAGAGAATCAAAGCTGGAGACACAGCTGTCCTTGAC
 TGTGAGGTCACTGGGGATCCCAAACCAAAAATATTTTGGTTGCTGCCTTCCAATGAC
 ATGATTTTCTTCTCCATTGATAGGTACACATTTTCATGCCAATGGGTCTTTGACCATC
 AACAAAGTGAACTGCTCGATTCTGGAGAGTACGTATGTGTAGCCCGAAATCCCAGT
 GGGGATGACACCAAAATGTACAACTGGATGTGGTCTCTAAACCTCCATTAATCAAT
 GGTCTGTATACAAACAGAACTGTTATTAAAGCCACAGCTGTGAGACATTCCAAAAAA
 CACTTTGACTGCAGAGCTGAAGGGACACCATCTCCTGAAGTCATGTGGATCATGCCA
 GACAATATTTTCTCACAGCCCCATACTATGGAAGCAGAATCACAGTCCATAAAAAT
 GGAACCTTGGAATTAGGAATGTGAGGCTTTCAGATTCAGCCGACTTTATCTGTGTG
 GCCCGAAATGAAGGTGGAGAGAGCGTGTTGGTAGTACAGTTAGAAGTACTGGAATG
 CTGAGAAGACCGACATTTAGAAATCCATTTAATGAAAAAATAGTTGCCCAGCTGGGA
 AAGTCCACAGCATTGAATTGCTCTGTTGATGGTAACCCACCACCTGAAATAATCTGG
 ATTTTACCAAATGGCACACGATTTTCCAATGGACCACAAAGTTATCAGTATCTGATA
 GCAAGCAATGGTTCTTTTATCATTTCTAAAACAACTCGGGAGGATGCAGGAAAATAT
 CGCTGTGCAGCTAGGAATAAAGTTGGCTATATTGAGAAATTAGTCATATTAGAAATT
 GGCCAGAAGCCAGTTATTCTTACCTATGCACCAGGGACAGTAAAAGGCATCAGTGGA
 GAATCTCTATCACTGCATTGTGTGTCTGATGGAATCCCTAAGCCAAATATCAAATGG
 ACTATGCCAAGTGGTTATGTAGTAGACAGGCCCTCAAATTAATGGGAAATACATATTG
 CATGACAATGGCACCTTAGTCATTAAAGAAGCAACAGCTTATGACAGAGGAACTAT
 ATCTGTAAGGCTCAAATAGTGTTGGTCATACACTGATTACTGTTCCAGTAATGATT
 GTAGCCTACCCTCCCCGAATTACAAATCGTCCACCCAGGAGTATTGTCACCAGGACA
 GGGGCAGCCTTTCAGCTCCACTGTGTGGCCTTGGGAGTTCCCAAGCCAGAAATCACA
 TGGGAGATGCCTGACCACTCCCTTCTCTCAACGGCAAGTAAAGAGAGGACACATGGA
 AGTGAGCAGCTTCACTTACAAGGTACCCTAGTCATTCAGAATCCCCAAACCTCCGAT
 TCTGGGATATACAAATGCACAGCAAAGAACCCACTTGGTAGTGATTATGCAGCAACC
 TATATTCAAGTAATCTGA

Figure 5b
Human OCP: predicted amino acid sequence

MKVKGRGITC LLVSFAVICL VATPGGKACP RRCACYMPTE VHCTFRYLTS
IPDSIPPNVE RINLGYNLSV RLMETDFSGL TKLELLMLHS NGIHTIPDKT
FSDLQALQVL KMSYNKVRKL QKDTFYGLRS LTRLHMDHNN IEFINPEVFY
GLNFLRLVHL EGNQLTKLHP DTFVSLSYLQ IFKISFIKFL YLSDNFLTSL
PQEMVSYMPD LDSLYLHGPN WTCDCHLKWL SDWIQEKPDV IKCKKDRSPS
SAQQCPLCMN PRTSKGKPLA MVSAAAFQCA KPTIDSSLKS KSLTILEDSS
SAFISPOGFM APFGSLTLNM TDQSGNEANM VCSIQKPSRT SPIAFTEEND
YIVLNTSFST FLVCNIDYGH IQPVWQILAL YSDSPLILER SHLLSETPQL
YYKYKQVAPK PEDIFTNIEA DLRADPSWLM QDQISLQLNR TATTFTSLQI
QYSSDAQITL PRAEMRPVKH KWTMISRDNN TKLEHTVLVG GTVGLNCPGQ
GDPTPHVDWL LADGSKVRAP YVSEDGRILI DKSGKLELQM ADSFDTGVYH
CISSNYDDAD ILTYRITVVE PLVEAYQENG IHHTVFIGET LDLPCHSTGI
PDASISWVIP GNNVLYQSSR DKKVLNNGTL RILQVTPKDQ GYYRCVAANP
SGVDFLIFQV SVKMKGQRPL EHDGETEGSG LDESNPIAHL KEPPGAQLRT
SALMEAEVGK HTSSTSKRHN YRELTQLRRG DSTHRRFREN RRHFPPSARR
IDPQHWAAALL EKAKKNAMPD KRENTTVSPP PVVTQLPNIP GEEDDSSGML
ALHEEFMVPA TKALNLPART VTADSRTISD SPMTNINYGT EFSPVVNSQI
LPPEEPTDFK LSTAIKTTAM SKNINPTMSS QIQGTTNQHS STVFPLLLGA
TEFQDSQMG RGREHFQSRP PITVRTMIKD VNVKMLSST NKLLLESVNT
TNSHQTSVRE VSEPRHNFY SHTTQILSTS TFPSPHTAA HSQFPIPRNS
TVNIPLFRRF GRQRKIGGRG RIISPYRTPV LRRHRYISFR STTRGSSEKS
TTAFSATVLN VTCLSCLPRE RLTTATAALS FPSAAPITFP KADIARVPSE
ESTTLVQNPL LLENKPSVE KTTPTIKYFR TEISQVTPTG AVMTYAPTISI
PMEKTHKVNA SYPRVSSNE AKRDSVITSS LSGAITKPPM TIIAITRFSR
RKIPWQQNFV NNHNPKGRLR NQHKVSLQKS TAVMLPKTSP ALPQRQSSPF
HFTTLSTSVM QIPSNLTLLTA HHTTTKTHNP GSLPTKKELP FPPLNPMPLPS
IISKDSSTKS IISTQTAIPA TTPTFPASVI TYETQTERSAR AQTIQREQEP
QKKNRTDPNI SPDQSSGFTT PTAMTPPALA FTHSPPENTT GISSTISFHS
RTLNLTDVIE ELAQASTQTL KSTIASSETTL SSKSHQSTTT RKASLDTPIP
PFLSSSATLM PVPISPPFTQ RAVTDTRGDS HFRMTNTVV KLHESSRHLN
QMPSSQLEPL TSSTSNNLLHS TMPALTTVK SQNSKLTPSP WAEYQFWHKP
YSDIAEKGGK PEVSMLATTG LSEATTLVSD WDGQKNTKKS DFDKKPVQEA
TTSKLLPFDS LSRYIFEKPR IVGGKAASFT IPANSDAFLP CEAVGNPLPT
IHWTRVSGLD LSRGNQNSRV QVLPNGTSLI QRVEIQDRGQ YLCSASNLFQ
TDHLHVTLSV VSYPPRILER RTKEITVHSG STVELKCRAE GRPSPTVTWI
LANQTVVSES SQGSRQAVVT VDGTLVLHNL SIYDRGFYKC VASNPGGQDS
LLVKIQVIAA PPVILEQRRQ VIVGTWGESL KLPCTAKGTP QPSVYVWLSD
GTEVKPLQFT NSKLFLFSNG TLYIRNLASS DRGTYECIAT SSTGSERRVV
MLTMEERVTS PRIEASQKR TEVNFGDKLL LNCSATGEPK PQIMWRLPSK
AVVDQWSWIH VYPNGSLFIG SVTEKDSGVY LCVARNKMGD DLILMHVSLR
LKPAKIDHKQ YFRKQVLHGK DFQVDCASG SPVPEISWSL PDGTMINNAM
QADDSGHRTR RYTLFNNGTL YFNKVGVAEE GDYTCYAQNT LGKDEMKVHL
TVITAAPRIR QSNKTNKRIK AGDTAVLDCE VTGDPKPKIF WLLPSNDMIS
FSIDRYTFHA NGSLTINKVK LLDSGEYVCV ARNPSGDDTK MYKLDVVSKEP
PLINGLYTNR TVIKATAVRH SKKHFDCAE GTPSPEVMWI MPDNIFLTAP
YYGSRITVHK NGTLEIRNVR LSDSADFICV ARNEGGSVL VVQLEVLEML
RRPTFRNPFN EKIVAQLGKS TALNCSVDGN PPPEIIWILP NGTRFSNGPQ
SYQYLIASNG SFIISKTTRE DAGKYRCAAR NKVGYIEKLV ILEIGQKPMI

[illegible]